OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 29, 2005, 11:18:45 ; Search time 11852 Seconds
                       (without alignments)
                       11328.839 Million cell updates/sec

Title: US-10-734-801-18
Perfect score: 2771
Sequence: 1 ggatccgcgtggaatgtgt.........atgggaccacacaaatctctc 2771
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 4708233 seqs, 2422767955 residues
Total number of hits satisfying chosen parameters: 9416466
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
                 1: gb_ba:*
                 2: gb_htg:*
                 3: gb_in:*
                 4: gb_om:*
                 5: gb_ov:*
                 6: gb_pat:*
                 7: gb_ph:*
                 8: gb_pl:*
                 9: gb_pr:*
                 10: gb_ro:*
                 11: gb_sts:*
                 12: gb_sy:*
                 13: gb_un:*
                 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2771	100.0	2771	6	AR478128	AR478128 Sequence
2	2771	100.0	5860	6	AR478127	AR478127 Sequence
3	2674	96.5	2674	6	AR478129	AR478129 Sequence
4	2327	84.0	2327	14	AR478130	AR478130 Sequence
5	1992.2	71.9	958	12	AB119282	AB119282 Hepatitis
6	1773.2	64.0	917	12	AB037684	AB037684 Cloning v
7	1766	63.7	6192	12	AF416989	AF416989 Synthetic
8	1763.2	63.6	5064	12	AX339209	AX339209 Sequence
9	1763.2	63.6	5064	12	U47297	U47297 Cloning vec
10	1763.2	63.6	5256	12	AX339208	AX339208 Sequence
11	1763.2	63.6	5256	12	U47296	U47296 Cloning vec
12	1763.2	63.6	612	12	AF416988	AF416988 Synthetic
13	1763.2	63.6	7014	6	CQ803129	CQ803129 Sequence
14	1758.2	63.5	7788	6	AX528380	AX528380 Sequence
15	1747.2	63.1	7312	6	AX528381	AX528381 Sequence
16	1697.2	61.2	6320	12	AB038600	AB038600 Cloning v
17	1657.6	59.8	4818	6	AX339207	AX339207 Sequence
18	1657.6	59.8	4818	12	U47295	U47295 Cloning vec
19	1657.6	59.8	4955	12	AF264724	AF264724 Cloning V
20	1657.6	59.8	5010	6	AX339210	AX339210 Sequence
21	1657.6	59.8	5010	12	U47298	U47298 Cloning vec
22	1657.6	59.8	5468	6	CQ849529	CQ849529 Sequence
23	1657.6	59.8	5960	6	BD084716	BD084716 Human ner
24	1657.6	59.8	6069	12	AF515711	AF515711 Synthetic
25	1657.6	59.8	8307	6	CQ849530	CQ849530 Sequence
26	1657.6	59.8	10716	12	AY603760	AY603760 Cloning v
27	1657.6	59.8	10820	12	AY603761	AY603761 Cloning v
28	1657.6	59.8	10837	12	AY603762	AY603762 Cloning v
29	1657.6	59.8	21404	6	BD167287	BD167287 Gene enco
30	1657.6	59.8	21404	6	BD170020	BD170020 Aip-1/FLA
31	1656.6	59.8	8712	6	AX179546	AX179546 Sequence
32	1656.4	59.8	13654	6	AR492053	AR492053 Sequence
33	1655.8	59.8	4553	12	AF027126	AF027126 Eukaryoti
34	1655.8	59.8	4643	12	AF027127	AF027127 Eukaryoti
35	1655.8	59.8	4711	12	AF027128	AF027128 Eukaryoti
36	1655.8	59.8	5376	12	AF027129	AF027129 Eukaryot
37	1655.2	59.7	1686	6	AR266991	AR266991 Sequence
38	1655.2	59.7	4413	6	CQ803117	CQ803117 Sequence
39	1655.2	59.7	7209	6	CQ803118	CQ803118 Sequence
40	1655.2	59.7	14947	6	AR920047	AR920047 Sequence
41	1655	59.7	12850	6	AF262031	AF262031 Cloning v
42	1654.4	59.7	6087	12	AF262031	AF262031 Cloning v
43	1654.2	59.7	4100	12	U47122	U47122 Cloning vec
44	1654.2	59.7	14194	6	AX052541	AX052541 Sequence
45	1653.8	59.7	10869	12	AY603757	AY603757 Cloning v

ALIGNMENTS

RESULT 1
AR478128

LOCUS   AR478128   2771 bp   DNA   linear   PAT 14-MAY-2004
DEFINITION Sequence 18 from patent US 6699657.
ACCESSION   AR478128
VERSION   AR478128.1 GI:47236755
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2771)
  AUTHORS   King,R.W., Jeffries,M.W. and Pasquinelli,C.
  TITLE   In vitro system for replication of RNA-dependent RNA polymerase
            (RDRP) viruses
  JOURNAL   Patent: US 6699657-A 18 02-MAR-2004;
FEATURES   Location/Qualifiers
     source   1..2771
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   100.0%; Score 2771; DB 6; Length 2771;
Best Local Similarity 100.0%;
Matches 2771; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GGATCCGCGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGATCCGCGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 60

Qy   61 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120

Qy   121 CTCCCCAGCAGGCAGGAGGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTCCCCAGCAGGCAGGAGGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180

Qy   181 GCCCCTAACTCCCGCCCATCCCGCCCCTAACTCCCGCCCAGTTCCGCCCCATTCTCCGCCCCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GCCCCTAACTCCCGCCCATCCCGCCCCTAACTCCCGCCCAGTTCCGCCCCATTCTCCGCCCCA 240

Qy   241 TGGCTGACTAATTTTTTTTATTTTATTATGCAGAGGCCGAGGCCGAGGCGGCTCTCGAGCTATT 300

```
Db  241  TGGCTGACTAATTTTTTTTTATTTATGCAGAGCGAGGCCGAGGGCCCTTCTGAGCTATT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  301  CCAGAAGTAGTGAGGAGGCCTTTTTTTGGGAGGCCTAGGGCTTTTGCAAAAAGCTTACATGATC  360
         ||||||||||||||||||  |||||||||||||
Db  301  CCAGAAGTAGTGAGGAGGCCTTTTTTTGGGAGGCCTAGGGCCTAGGGCCTTACATGATC  360

Qy  361  TGCAGAGAGGCCACTAATCAGCACTCTCTCGCCAGTCATGCCGGCTCACGGGACCTTTCACAGCT  420
Db  361  TGCAGAGAGGCCACTATCAGCACTCTCTCGCCAGTCATGCCGGCTCACGGGACCTTTCACAGCT  420

Qy  421  AGCCGTGACTAGGGCTAAGATGGGAGGCCACCCATTAAAGGAAGGAAGGAAGAAAAGGAAAA  480
Db  421  AGCCGTGACTAGGGCTAAGATGGGAGGCCACCCATTAAAGGAAGGAAGGAAGAAAAGGAAAA  480

... (DNA sequence alignment continues)

Db  1321 AGGGGAGCGGCCACCAGGAAGCAAATTTCGTGTTGTAAATTAGATAAATCGTATTTGTCAATCAGA  1380

Qy  2401 TCGCGACCCAACACTACTCGGCTAGCAGTCTTGCGGGGGGCACGCGCCCAAAATCTCCAGGCAT  2460
Db  2401 TCGCGACCCAACACTACTCGGCTAGCAGTCTTGCGGGGGGCACGCGCCCAAAATCTCCAGGCAT  2460
```

```
RESULT 2
LOCUS       AR478127
LOCUS       AR478127          5860 bp    DNA        linear      PAT 14-MAY-2004
DEFINITION  Sequence 17 from patent US 6699657.
ACCESSION   AR478127
VERSION     AR478127.1    GI:47236754
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 5860)
  AUTHORS   King,R.W., Jeffries,M.W. and Pasquinelli,C.
  TITLE     In vitro system for replication of RNA-dependent RNA polymerase
            (RDRP) viruses
  JOURNAL   Patent: US 6699657-A 17 02-MAR-2004;
FEATURES             Location/Qualifiers
     source          1..5860
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match        100.0%;  Score 2771;  DB 6;  Length 5860;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 2771;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy       1 GGATCCGCTGTGGAAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGCCTCCCCAGCAGG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGATCCGCTGTGGAAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGCCTCCCCAGCAGG   60

Qy      61 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG  120

Qy     121 CTCCCCAGCCAGGCAGAAGTATGCAAAGCATCTCAATTAGTCAGCAACCAGGTGTCC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTCCCCAGCCAGGCAGAAGTATGCAAAGCATCTCAATTAGTCAGCAACCAGGTGTCC  180

Qy     181 GCCCCTAACTCCGCCCATCCCGCCCCACTCCGCCCATCCCGCCCATCCCGCCCCA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCCCCTAACTCCGCCCATCCCGCCCCACTCCGCCCATCCCGCCCATCCCGCCCCA  240

Qy     241 TGGCTGACTAATTTTTTTATTTTTATGCAGAGGGCCTGAGGCCTGAGCTTATT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGGCTGACTAATTTTTTTATTTTTATGCAGAGGGCCTGAGGCCTGAGCTTATT  300

Qy     301 CCAGAAGTAGTGAGGAGGAGCCTTTTTGGAGGAGGAATAGGGTTGGCACCAGCAGC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CCAGAAGTAGTGAGGAGGAGCCTTTTTGGAGGAGGAATAGGGTTGGCACCAGCAGC  360
```

```
Db     301 CCAGAAGTAGTGAGGAGGAGCCTTTTTGGAGGAGGAATAGGGTTGGCACCAGCAGC  360
Db     361 TGCAGAGAGGCCAGTATACAGCACTCTTCTCGAGTCATGCGGGCTCACGGACCTTTCACAGCT  420
Db     361 TGCAGAGAGGCCAGTATACAGCACTCTTCTCGAGTCATGCGGGCTCACGGACCTTTCACAGCT  420

Qy     421 AGCCGTGACTAGGGCTAAGATGGAGGCCACCATTAAAGAAGGAAAGGAAAGGAAAA  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AGCCGTGACTAGGGCTAAGATGGAGGCCACCATTAAAGAAGGAAAGGAAAGGAAAA  480

Qy     481 AGAAGGAAAGAAGAAA...AGGAAAGGAAAGGAAAAAAAAAAAAAAAAAAAAAAA  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AGAAGGAAAGAAGAAA...AGGAAAGGAAAGGAAAAAAAAAAAAAAAAAAAAAAA  540

Qy     541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  600

Qy     601 GGCCGGAGTGTTTACCCCAAACCTTTAAACGGCCGATCTTTCCGCCCCTTTTATG  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GGCCGGAGTGTTTACCCCAAACCTTTAAACGGCCGATCTTTCCGCCCCTTTTATG  660

Qy     661 AGGATCTCTCTGATTTTTCTTGCGTCGGAGTTTTCGGTAAGACCTTTCGGGTACTTCGTCC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AGGATCTCTCTGATTTTTCTTGCGTCGGAGTTTTCGGTAAGACCTTTCGGGTACTTCGTCC  720

Qy     721 ACAAACACAAACTCCTCCCGCGCAACTGTTTTTGGTTGTTGTTACTTGACTGGCGGTAATCC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACAAACACAAACTCCTCCCGCGCAACTGTTTTTGGTTGTTGTTACTTGACTGGCGGTAATCC  780

Qy     781 ACGAATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCTTCCAAAACAACAACGCGGCGGGAAGT  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ACGAATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCTTCCAAAACAACAACGCGGCGGGAAGT  840

Qy     841 TCACCGGCGGTCATCGTCGGGGAAGACCTGGCACCTGCCGTCGGCTCGGAAGATGTTGGGG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCACCGGCGGTCATCGTCGGGGAAGACCTGGCACCTGCCGTCGGCTCGGAAGATGTTGGGG  900

Qy     901 AGCAAGATGGATTCCAATTCAGCGGGAGCGCACCTGATAGCCTTTGTACTTAATCAGAGAC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 AGCAAGATGGATTCCAATTCAGCGGGAGCGCACCTGATAGCCTTTGTACTTAATCAGAGAC  960

Qy     961 TTCAGGCCGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCCAGTAAGCTATGTCTCCAGAA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTCAGGCCGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCCAGTAAGCTATGTCTCCAGAA 1020

Qy    1021 TGTAGCCATGCCATTCCTTGTCAATCAAGGCGGTGGTCGCTTCCGGATTGTTTACATAACG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TGTAGCCATGCCATTCCTTGTCAATCAAGGCGGTGGTCGCTTCCGGATTGTTTACATAACG 1080

Qy    1081 GACATAATCATTAGGAACCTCTCACACAGTTCCGCCTCTTTGATTAACGCCCACGGTTTTC 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GACATAATCATTAGGAACCTCTCACACAGTTCCGCCTCTTTGATTAACGCCCACGGTTTTC 1140

Qy    1141 CCGGTATCCAGATCCACAACACCTTCGCTTCAAAAATGGAACAACTTTACCGGACCGCGCC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CCGGTATCCAGATCCACAACACCTTCGCTTCAAAAATGGAACAACTTTACCGGACCGCGCC 1200

Qy    1201 GGTTTATCATCCCCCTCGGGGTGTAAATCAGAAATAGCTGATGTAGTCAGTGAGCCCATAT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GGTTTATCATCCCCCTCGGGGTGTAAATCAGAAATAGCTGATGTAGTCAGTGAGCCCATAT 1260

Qy    1261 CCTTGCCTGATACCTGGGCAGAATGGAAACCTTCTTGGCACCGCTTCCCCGGACTTCCTTAGAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CCTTGCCTGATACCTGGGCAGAATGGAAACCTTCTTGGCACCGCTTCCCCGGACTTCCTTAGAG 1320

Qy    1321 AGGGGAGCGCCACCAGAAGCAAATTTCGTGTAAATTAGATAAAATCGTATTTGTCAATCAGA 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 AGGGGAGCGCCACCAGAAGCAAATTTCGTGTAAATTAGATAAAATCGTATTTGTCAATCAGA 1380

Qy    1381 GTGCTTTTGGCGGAAGAAGGAAGGAGAATAGGGTTGGCACCAGCAGCGCCACTTTGAATCTTGTAA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GTGCTTTTGGCGGAAGAAGGAAGGAGAATAGGGTTGGCACCAGCAGCGCCACTTTGAATCTTGTAA 1440
```

```
Qy    2461 TGAGCGGGGGGTTATTCCAAGAAAAGGACCCCGGTCGTCCTGGCCAATTCCGTGTACTCACCCGGT 2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 TGAGCGGGGGGTTATTCCAAGAAAAGGACCCCGGTCGTCCTGGCCAATTCCGTGTACTCACCCGGT 2520

Qy    2521 TCCGCAGACCACTATGGGCTCTCCCGGGGAGGGGCTGCACGACACTCATA 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 TCCGCAGACCACTATGGGCTCTCCCGGGGAGGGGCTGCACGACACTCATA 2580

Qy    2581 CTPAACGCCATGGCCTAGACGGTTTTCTTGGTGTGAGACAGAGTTGGATT 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 CTPAACGCCATGGCCTAGACGGTTTTCTTGGTGTGAGACAGAGTTGGATT 2640

Qy    2641 CATGGTGGAGTGTTCGCCCCATCAGGGGGTCCCACTCCTCCTCGC 2700
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 CATGGTGGAGTGTTCGCCCCATCAGGGGGTCCCACTCCTCCTCGC 2700

Qy    2701 TGGCCGCGGCTGGGCAACATTCCGAGGGGGAACCGGTTAATGGGCGAATGGGACCC 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 TGGCCGCGGCTGGGCAACATTCCGAGGGGGAACCGGTTAATGGGCGAATGGGACCC 2760

Qy    2761 ACAAATCTCTC 2771
           |||||||||||
Db    2761 ACAAATCTCTC 2771
```

```
Qy   1441 TCCTGAAGGCTTCCTCCAGAAACAGCTCTTCTTCTTCAAATCTCATACATTCATTAAGACGACTCGAAAT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCCTGAAGGCTTCCTCCAGAAACAGCTCTTCTTCTTCAAATCTCATACATTCATTAAGACGACTCGAAAT 1500

Qy   1501 CCACATATCAAAATATCCGAGTGTAGTAAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CCACATATCAAAATATCCGAGTGTAGTAAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560

Qy   1561 CTTAAAAATCGCAGTATCCGAAATGATTTGATTGCCAAAAATAGAGATCTCTTGGCATGCGAG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CTTAAAAATCGCAGTATCCGAAATGATTTGATTGCCAAAAATAGAGATCTCTTGGCATGCGAG 1620

Qy   1621 AATCTCACCGCAGGCAGTTCTTATGAGGCAGCAGCGCACCTTTAGGCAGACCAGTAGATCCA 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AATCTCACCGCAGGCAGTTCTTATGAGGCAGCAGCGCACCTTTAGGCAGACCAGTAGATCCA 1680

Qy   1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAAGGACTCTGGCACAAAATCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAAGGACTCTGGCACAAAATCG 1740

Qy   1741 TATTCATTAAAAACCGGGAGGTAGATGATGAGATGTGACGAAGCGTGTACATCGACTGAAAATCCC 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TATTCATTAAAAACCGGGAGGTAGATGATGAGATGTGACGAAGCGTGTACATCGACTGAAAATCCC 1800

Qy   1801 TGGGTAATCCGTTTTAGAATCCATGAATCGATGAATAATATTTTTTGGATGAATTGGGAGCTTTTTTTGC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGGGTAATCCGTTTTAGAATCCATGAATCGATGAATAATATTTTTTGGATGAATTGGGAGCTTTTTTTGC 1860

Qy   1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACCGGGTAGGCTGGCAAATG 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACCGGGTAGGCTGGCAAATG 1920

Qy   1921 CCCATACTGTTGAGCAATTCACGTTCATTATTCATTAGTCGTTCGTCGGGCAAACTGCAACT 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CCCATACTGTTGAGCAATTCACGTTCATTATTCATTAGTCGTTCGTCGGGCAAACTGCAACT 1980

Qy   1981 CCGATAAAATAACGCCCAACACCGGCATAAAGAAGAATTGAAGAGAGTTTTCACTGCATACG 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CCGATAAAATAACGCCCAACACCGGCATAAAGAAGAATTGAAGAGAGTTTTCACTGCATACG 2040

Qy   2041 ACGATTCTGATTTGTATTCAGCGCCATATCGTTTTCATAGCTTCTGCCAACCGAACGGAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACGATTCTGATTTGTATTCAGCGCCATATCGTTTTCATAGCTTCTGCCAACCGAACGGAC 2100

Qy   2101 ATTTCGAAGTACTCAGCGTAAGTGAATGTGCATCTGTAAAAGCAATT 2160
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATTTCGAAGTACTCAGCGTAAGTGAATGTGCATCTGTAAAAGCAATT 2160

Qy   2161 GTTCCAGGAACCAGGGGCGTATCTCTTCAGTAGCCTTATGCAGTTCCA 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GTTCCAGGAACCAGGGGCGTATCTCTTCAGTAGCCTTATGCAGTTCCA 2220

Qy   2221 TCTTCCCAGCGGGATAGAAATGGCGCGCGGGCCTTTCTTTATGTTTTTTGGGCGTTCTTCCATGGGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TCTTCCCAGCGGGATAGAAATGGCGCGCGGGCCTTTCTTTATGTTTTTTGGGCGTTCTTCCATGGGA 2280

Qy   2281 CGTCGGGTTGGTGTTACGTTTGGTTTTTCTTGAGGTTTAGGATTCGTGTCATGATGATGCAC 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CGTCGGGTTGGTGTTACGTTTGGTTTTTCTTGAGGTTTAGGATTCGTGTCATGATGATGCAC 2340

Qy   2341 GGTCTACGGAGACCTCCCGGGGGGCACTCGCAAGCACACCTATCAGGCAGTACACCAAGGCCTT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GGTCTACGGAGACCTCCCGGGGGGCACTCGCAAGCACACCTATCAGGCAGTACACCAAGGCCTT 2400

Qy   2401 TCGCCGACCCAAACACTACTCGGCTACCAGTCTTGCGGGCGCACGCCCAAAATCTCCAGGCAT 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TCGCCGACCCAAACACTACTCGGCTACCAGTCTTGCGGGCGCACGCCCAAAATCTCCAGGCAT 2460

Qy   2461 TGAGCGGGGGTTATCCAAGAAAAAGGACCCGGTCGTCCTGGCAAATTCGGTGTAGTACTCACCGGT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TGAGCGGGGGTTATCCAAGAAAAAGGACCCGGTCGTCCTGGCAAATTCGGTGTAGTACTCACCGGT 2520

Qy   2521 TCCGCAGACCACTATGGCTCTCCCGGGGGTCTCCGGGGAGGGGGTCCTTGCACGACACTCATA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TCCGCAGACCACTATGGCTCTCCCGGGGGTCTCCGGGGAGGGGGTCCTTGCACGACACTCATA 2580

Qy   2581 CTAACGCCCATGGCTAGACGCTTTTCTGCGTGTCGAAAGACAGTAGTTCCTCACAGGGGAGTTGATT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CTAACGCCCATGGCTAGACGCTTTTCTGCGTGTCGAAAGACAGTAGTTCCTCACAGGGGAGTTGATT 2640

Qy   2641 CATGGTGGGAGGTGTCGCCCCCCATCAGGGGGCGTGGGGGGCTGGGGCCATGGTCCCCAGCCTCTCGC 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 CATGGTGGGAGGTGTCGCCCCCCATCAGGGGGCGTGGGGGGCTGGGGCCATGGTCCCCAGCCTCTCGC 2700

Qy   2701 TGGCGCCGGGGCTGGGGCAAACATTCCGAGGGAGGGGACCGTCCCCTCGGTAATGGGGAAATGGGACC 2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TGGCGCCGGGGCTGGGGCAAACATTCCGAGGGAGGGGACCGTCCCCTCGGTAATGGGGAAATGGGACC 2760

Qy   2761 ACAAATCTCTC 2771
          |||||||||||
Db   2761 ACAAATCTCTC 2771
```

RESULT 3

```
LOCUS       AR478129                2674 bp    DNA     linear   PAT 14-MAY-2004
DEFINITION  Sequence 19 from patent US 6699657.
ACCESSION   AR478129
VERSION     AR478129.1  GI:47236756
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2674)
  AUTHORS   King,R.W., Jeffries,M.W. and Pasquinelli,C.
  TITLE     In vitro system for replication of RNA-dependent RNA polymerase
            (RDRP) viruses
  JOURNAL   Patent: US 6699657-A 19 02-MAR-2004;
FEATURES             Location/Qualifiers
     source          1..2674
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match    96.5%;  Score 2674;  DB 6;  Length 2674;
Best Local Similarity  100.0%;  Pred. No. 0;  Mismatches 0;  Indels 0;  Gaps 0;
Matches 2674;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy      1 GGATCCGCTCTGGGAATGTGTGTCAGTTAGGGTGTGGAAAGCTCCCCAGGCGCTCCCAGCGCAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGATCCGCTCTGGGAATGTGTGTCAGTTAGGGTGTGGAAAGCTCCCCAGGCGCTCCCAGCGCAGG 60

Qy     61 CAGAGAAGTATGCAAAGCATGCCATCTCAAATTAGTCAGCAACAACCAGGTGTTGGAAAGTCCCCAGG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGAGAAGTATGCAAAGCATGCCATCTCAAATTAGTCAGCAACAACCAGGTGTTGGAAAGTCCCCAGG 120

Qy    121 CTCCCCAGCAGGCAGAGAAGTATGCAAAGCATGCATCTCAAATTAGTCAGCAACCAGTAGTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCCCCAGCAGGCAGAGAAGTATGCAAAGCATGCATCTCAAATTAGTCAGCAACCAGTAGTCCC 180

Qy    181 GCCCCTAACTCCGCCCATCCCGCCCCCAGTTCCGCGCCCATTCTCGCCGCCCCCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCCCCTAACTCCGCCCATCCCGCCCCCAGTTCCGCGCCCATTCTCGCCGCCCCCA 240

Qy    241 TGGCTGACTAATTTTTTTATTATTGCAGAGGCCGAGGCGCCCTCTCGAGCTATT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGCTGACTAATTTTTTTATTATTGCAGAGGCCGAGGCGCCCTCTCGAGCTATT 300

Qy    301 CCAGAAGTAGTGAGGAGGCTTTTTTGGGAGGCCTAGGCTTTTGCAAAAGGTTACATGATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCAGAAGTAGTGAGGAGGCTTTTTTGGGAGGCCTAGGCTTTTGCAAAAGGTTACATGATC 360

Qy    361 TGCAGAGAGGCCCAGTATCAGCACTCTCTGCAGTCATGCGCGCTCACGGACCTTTTCACAGCT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGCAGAGAGGCCCAGTATCAGCACTCTCTGCAGTCATGCGCGCTCACGGACCTTTTCACAGCT 420
```

```

```
Qy   421 AGCCGTGACTAGGGCTAAGATGGAGCCACCATTAAAGAAGGAAGGAAAAGAAAGGAAAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCCGTGACTAGGGCTAAGATGGAGCCACCATTAAAGAAGGAAGGAAAAGAAAGGAAAAA 480

Qy   481 AGAAGGAAAGAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAGAAAAAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGAAGGAAAGAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAGAAAAAAAAAA 540

Qy   541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAATGGCCTAAGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAATGGCCTAAGA 600

Qy   601 GGCCGGAGTGTTTACCCCAACCTTTAAACGGCGATCTTTCCGCCCTTCTTGGCCTTTATG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCCGGAGTGTTTACCCCAACCTTTAAACGGCGATCTTTCCGCCCTTCTTGGCCTTTATG 660

Qy   661 AGGATCTCTCTGATTTTTCTTGCGTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AGGATCTCTCTGATTTTTCTTGCGTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCC 720

Qy   721 ACAAACACAACTCCTCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACAAACACAACTCCTCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCC 780

Qy   781 ACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACAACAACGGCGGCGGGAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACAACAACGGCGGCGGGAAGT 840

Qy   841 TCACCGGCGTCATCGTCGGGAAGACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCACCGGCGTCATCGTCGGGAAGACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGG 900

Qy   901 AGCAAGATGGATTCCAATTCAGCGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AGCAAGATGGATTCCAATTCAGCGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGAC 960

Qy   961 TTCAGGCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTCAGGCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAA 1020

Qy   1021 TGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCG 1080

Qy   1081 GACATAATCATAGGACCTCTCACACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GACATAATCATAGGACCTCTCACACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTC 1140

Qy   1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCC 1200

Qy   1201 GGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATAT 1260

Qy   1261 CCTTGCCTGATACCTGGCAGATGGAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CCTTGCCTGATACCTGGCAGATGGAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAG 1320

Qy   1321 AGGGGAGCGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGGAGCGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGA 1380

Qy   1381 GTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAA 1440

Qy   1441 TCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAAT 1500
```

```
Qy       1501 CCACATATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCACATATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560

Qy       1561 CTTAAAATCGCAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTTAAAATCGCAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAG 1620

Qy       1621 AATCTCACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATCTCACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCA 1680

Qy       1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCG 1740

Qy       1741 TATTCATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TATTCATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCC 1800

Qy       1801 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGC 1860

Qy       1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATG 1920

Qy       1921 CCCATACTGTTGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CCCATACTGTTGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACT 1980

Qy       1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040

Qy       2041 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGAC 2100

Qy       2101 ATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATT 2160

Qy       2161 GTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2220

Qy       2221 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGGGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGGGA 2280

Qy       2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTGCTCATGATGCAC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTGCTCATGATGCAC 2340

Qy       2341 GGTCTACGAGACCTCCCGGGGCACTCGCAAGCACCCTATCAGGCAGTACCACAAGGCCTT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GGTCTACGAGACCTCCCGGGGCACTCGCAAGCACCCTATCAGGCAGTACCACAAGGCCTT 2400

Qy       2401 TCGCGACCCAACACTACTCGGCTAGCAGTCTTGCGGGGGCACGCCCAAATCTCCAGGCAT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCGCGACCCAACACTACTCGGCTAGCAGTCTTGCGGGGGCACGCCCAAATCTCCAGGCAT 2460

Qy       2461 TGAGCGGGGTTATCCAAGAAAGGACCCGGTCGTCCTGGCAATTCCGGTGTACTCACCGGT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TGAGCGGGGTTATCCAAGAAAGGACCCGGTCGTCCTGGCAATTCCGGTGTACTCACCGGT 2520

Qy       2521 TCCGCAGACCACTATGGCTCTCCCGGGAGGGGGGGTCCTGGAGGCTGCACGACACTCATA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCCGCAGACCACTATGGCTCTCCCGGGAGGGGGGGTCCTGGAGGCTGCACGACACTCATA 2580

Qy       2581 CTAACGCCATGGCTAGACGCTTTCTGCGTGAAGACAGTAGTTCCTCACAGGGGAGTGATT 2640
```

RESULT 4

Qy   2641 CATGGTGGGACTGTGTCGCCCCCATCAGGGGGGGCTGGC 2674
             ||||||||||||||||||||||||||||||||||||||||
Db   2641 CATGGTGGGACTGTGTCGCCCCCATCAGGGGGGGCTGGC 2674

Qy   2641 CATGGTGGGACTGTGTCGCCCCCATCAGGGGGGGCTGGC 2674
             ||||||||||||||||||||||||||||||||||||||||
Db   2641 CATGGTGGGACTGTGTCGCCCCCATCAGGGGGGGCTGGC 2674

Db   2581 CTAACGCCATGGCTAGACGCCGTTTCTCCGTGAAGACAGTAGTTCCTCACAGGGAGAGTGGATT 2640
```

LOCUS   AR478130   2327 bp   DNA   linear   PAT 14-MAY-2004
DEFINITION Sequence 20 from patent US 6699657.
ACCESSION   AR478130
VERSION   AR478130.1 GI:47236757
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2327)
  AUTHORS   King,R.W., Jeffries,M.W. and Pasquinelli,C.
  TITLE   In vitro system for replication of RNA-dependent RNA polymerase
            (RDRP) viruses
  JOURNAL   Patent: US 6699657-A 20 02-MAR-2004;
FEATURES   Location/Qualifiers
     source   1..2327
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   84.0%; Score 2327; DB 6; Length 2327;
Best Local Similarity   100.0%;
Matches 2327; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy       348 AGCTTACATGATCTGCAGAGAGGCCAGTAATCAGCACTCACTCTTCTGCAGTCATGCGGCTCACGG  407
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AGCTTACATGATCTGCAGAGAGGCCAGTAATCAGCACTCACTCTTCTGCAGTCATGCGGCTCACGG   60

Qy       408 ACCTTTCACAGCTAGCCGTGCTAAGATGGAGCCACCAATTAAAGAGAAGGAAGGAA  467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ACCTTTCACAGCTAGCCGTGCTAAGATGGAGCCACCAATTAAAGAGAAGGAAGGAA  120

Qy       468 AAGGAAAGGAAAAAGAAGAAGAAGAAAAAAAAAAAAAAAAAAAAAAAA  527
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AAGGAAAGGAAAAAGAAGAAGAAGAAAAAAAAAAAAAAAAAAAAAAAA  180

Qy       528 AAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGG  587
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGG  240

Qy       588 AAAATGGGCTAAGAGAGGGCCGAGTGTTTACCCCACCACTCTTTAAACGGGGA  647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 AAAATGGGCTAAGAGAGGGCCGAGTGTTTACCCCACCACTCTTTAAACGGGGA  300

Qy       648 CTTGGCCCTTTATGAGGATCTCTCGATTTTTTTCTTGCGTGTAAGACCTTT  707
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CTTGGCCCTTTATGAGGATCTCTCGATTTTTTTCTTGCGTGTAAGACCTTT  360

Qy       708 CGGTTACTTCGTCCACAAACACAACTCCCTCCGCGCAACTTCGGTTTTGGTTT  767
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CGGTTACTTCGTCCACAAACACAACTCCCTCCGCGCAACTTCGGTTTTGGTTT  420

Qy       768 GGGGACGTAATCCACGAGATCTCTTTTTTCCGTCATCGTCTTTTCCGTCTTCCGTT  827
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GGGGACGTAATCCACGAGATCTCTTTTTTCCGTCATCGTCTTTTCCGTCTTCCGTT  480

Qy       828 GGGCGGCGGGGAAAGTTCACCGCCGTCGGGGAGAGACCTGCGTGCTGGAAGAT  887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GGGCGGCGGGGAAAGTTCACCGCCGTCGGGGAGAGACCTGCGTGCTGGAAGAT  540

Qy       888 GTTGGGGTGTTGGAGCAAGATGGAATTCCAATTCAGCGGGGAGCGAGCACCTGTAT  947
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GTTGGGGTGTTGGAGCAAGATGGAATTCCAATTCAGCGGGGAGCGAGCACCTGTA  600

Qy       948 CTTAATCAGAGACTTCAGGGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGC 1007
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CTTAATCAGAGACTTCAGGGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGC  660

Qy      1008 TATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGTCTTCCGGATT 1067
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 TATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGTCTTCCGGATT  720

Qy      1068 GTTTACATTACACCGGACACATATCATTAGGACCTCTCACACAGTTGCCTCTTTGATTAAC 1127
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GTTTACATTACACCGGACACATATCATTAGGACCTCTCACACAGTTGCCTCTTTGATTAAC  780

Qy      1128 GCCCAGCGGTTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAAATGGAACAACTTT 1187
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GCCCAGCGGTTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAAATGGAACAACTTT  840

Qy      1188 ACCGACCGGCGCGGGGTTTATCATCCCCCTCGGGTTGTAATCAGAATAGCTGATTGTAGTCTC 1247
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 ACCGACCGGCGCGGGGTTTATCATCCCCCTCGGGTTGTAATCAGAATAGCTGATTGTAGTCTC  900

Qy      1248 AGTGAGCCCCATATCCTTGCCTGATACCTGCGATGGAACCTCTGGGTGTAATCAGAATCAGAGT 1307
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 AGTGAGCCCCATATCCTTGCCTGATACCTGCGATGGAACCTCTGGGTGTAATCAGAATCAGAGT  960

Qy      1308 GACTTCCTTAGAGAGGGGAGCGCGCACCAGAAGCAAATTTCGTGTAAAATTAGATAAAATCGTA 1367
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GACTTCCTTAGAGAGGGGAGCGCGCACCAGAAGCAAATTTCGTGTAAAATTAGATAAAATCGTA 1020

Qy      1368 TTTGTCAATCAGAGTGCTTTTGGGGAAGAGGAGAAGAATAGGGGTTGCACCAGCAGCCACT 1427
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 TTTGTCAATCAGAGTGCTTTTGGGGAAGAGGAGAAGAATAGGGGTTGCACCAGCAGCCACT 1080

Qy      1428 TTGAATCTTGATTCCTTGAATCCTGAAGGGCTCCTCAGAAAACAGCTCTTCTTCTTCAGCATTAA 1487
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TTGAATCTTGATTCCTTGAATCCTGAAGGGCTCCTCAGAAAACAGCTCTTCTTCTTCAGCATTAA 1140

Qy      1488 GACGACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAAACATTCCAAAAAACCGTGATG 1547
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GACGACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAAACATTCCAAAAAACCGTGATG 1200

Qy      1548 GAATGGAACAACACTTAAAATCGCAGTATCCGGAATGAATGATTTGATTGACGATAGGATC 1607
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GAATGGAACAACACTTAAAATCGCAGTATCCGGAATGAATGATTTGATTGACGATAGGATC 1260

Qy      1608 TCTGGCCATGCCGAGAATCTCACGCAGCCAGTTCTTATGAGGCAGCACCTTTAGGCAG 1667
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 TCTGGCCATGCCGAGAATCTCACGCAGCCAGTTCTTATGAGGCAGCACCTTTAGGCAG 1320

Qy      1668 ACCAGTAGATCCAGAGGAGTTCATGTCCATCGTCATCGTCCCTATCGAACGGACTC 1727
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ACCAGTAGATCCAGAGGAGTTCATGTCCATCGTCATCGTCCCTATCGAACGGACTC 1380

Qy      1728 TGGCACAAAATCGTATTCATTAAAACCGGAGGTGAGATGAGTGACGGAACGTGTACAT 1787
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 TGGCACAAAATCGTATTCATTAAAACCGGAGGTGAGATGAGTGACGGAACGTGTACAT 1440

Qy      1788 CGACTGAAAATCCCGTTTAGAAATCCATGAAAATCATGAAAATTATTTTTTTGGATGATTGG 1847
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 CGACTGAAAATCCCGTTTAGAAATCCATGAAAATCATGAAAATTATTTTTTTGGATGATTGG 1500

Qy      1848 GAGCTTTTTTTTGCACGTTCAAAAATTTTTTTGCAACCCCTTTTGGAAAACGAACACCGGT 1907
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GAGCTTTTTTTTGCACGTTCAAAAATTTTTTTGCAACCCCTTTTGGAAAACGAACACCGGT 1560

Qy      1908 AGGCTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTAATAAATGTCGTTCGCGGG 1967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 AGGCTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTAATAAATGTCGTTCGCGGG 1620

Qy      1968 CGCAACTGCAAACTCCGATAAAATAACGCGCCCACACGGCATAAAGAAATTGAACGGAGTT 2027
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 CGCAACTGCAAACTCCGATAAAATAACGCGCCCACACGGCATAAAGAAATTGAACGGAGTT 1680

Qy      2028 TTCACTGCATACGACGAATCTCTGTGATTGTATTCAGCCCCATATCGTTCATAGCTTCTGC 2087
Db           ...
```

```
Db   1681  TTCACTGCATACGACGAGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGC  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2088  CAACCGAAACGACACATTTCGAAGTACTCACGCGTAAGTAGATGTCCACCTTCGATATGTGCATC  2147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1741  CAACCGAACGGACACATTTCGAAGTACTCACGGTAAGTGATGTCCACCTTCGATATGTGCATC  1800

Qy   2148  TGTAAAAGCAAATTGTTCCAGGAACCAGGGCGTATCTTCATAGCCTTATGCAGTTGGTC  2207
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  TGTAAAAGCAAATTGTTCCAGGAACCAGGGCGTATCTTCATAGCCTTATGCAGTTGGTC  1860

Db   1861  TCCAGCGCGTTCCATCTTCCAGCGGATAGAAATGGCGCGCCCGGGCCCTTCTTCTTATGTTTTTTGGC  1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2268  GTCTTCCATGGGACGTCGGTTGGTGTTTGGTGTTTGGTTTGGAGGTTTAGGAGAATTTCGT  2327
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1921  GTCTTCCATGGGACGTCGGTTGGTGTTGTTTCTTTGGTTTGGAGGTTTAGGAGATTTGT  1980

Qy   2208  TCCAGCGCGGTTCCATCTTCCAGCGGATAGAAATGGCGCGCCCGGGCCCTTCTTCTTATGTTTTTTGGC  2267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  GCTCATGATGCACGGTCTACGAGACCTCCGGGGCCGAAGCCAGCACCCTATCAGGCCAGT  2040

Qy   2328  GCTCATGATGCACGGTCTACGAGACCTCCGGGGCCGAAGCCAGCACCCTATCAGGCCAGT  2387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  ACCACAAGGCCTTTCGGACCCAAACAACTACCTCCCGGGCCCACTCGCAAGCCACCTATCAGGCCCA  2100

Qy   2388  ACCACAAGGCCTTTCGGCACCCAAACAACTACTCGGCTACCAGTCTTGCGGGGGCCCACGCCCA  2447
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  AATCTCCAGGCAATTGAGCGGGGGTTATTCCAAGAAAAGGACCCCGGTCGTCCTGGCAATTCCGG  2160

Qy   2448  AATCTCCAGGCAATTGAGCGGGGGTTATCCAAGAAAAGGACCCCGGTCGTCCTGGCAATTCCGG  2507
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  TGTACTCACCCGGTTCCGCGAGACCACTATGGCTCTCCCGGGGAGGGGGTCCTGGAGGCTG  2220

Qy   2508  TGTACTCACCCGGTTCCGCGAGACCACTATGGCTCTCCCGGGGAGGGGGTCCTGGAGGCTG  2567
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  CACGACACTCATACTAACGACCATGGCTCTTTCTGCGGGTCAGACAGTAGTTCCTCA  2280

Qy   2628  CAGGGGAGTGATTCATCGGTGGTGGTGCGCCCCCCATCAGGGGGCGTGGC  2674
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  CAGGGGAGTGATTCATCGGTGGTGGTGCGCCCCCCATCAGGGGGCGTGGC  2327

Qy   2568  CACGACACTCATACTAACGACCATGGCTCTTTCTGCGGGTCAGACAGTAGTTCCTCA  2627
```

RESULT 5
AB119282/c

LOCUS   AB119282   9658 bp   RNA   linear   VRL 27-MAR-2004
DEFINITION Hepatitis C virus gene for fusion protein, Feo, complete cds.
ACCESSION   AB119282
VERSION   AB119282.1 GI:37991671
KEYWORDS   .
SOURCE   Hepatitis C virus
  ORGANISM Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.

REFERENCE   1
  AUTHORS   Yokota,T., Sakamoto,N., Enomoto,N., Tanabe,Y., Miyagishi,M.,
            Maekawa,S., Yi,L., Kurosaki,M., Taira,K., Watanabe,M. and
            Mizuawa,H.
  TITLE   Inhibition of intracellular hepatitis C virus replication by
            synthetic and vector-derived small interfering RNAs
  JOURNAL   EMBO Rep. 4 (6), 602-608 (2003)
  MEDLINE   22625416
   PUBMED   12740604
REFERENCE   2
  AUTHORS   Tanabe,Y., Sakamoto,N., Enomoto,N., Kurosaki,M., Ueda,E.,
            Maekawa,S., Yamashiro,T., Nakagawa,M., Chen,C.H., Kanazawa,N.,
            Kakinuma,S. and Watanabe,M.
  TITLE   Synergistic Inhibition of intracellular Hepatitis C virus
            Replication by Combination of Ribavirin and Interferon-alpha
  JOURNAL   J. Infect. Dis. 189 (7), 1129-1139 (2004)
   PUBMED   15031779
REFERENCE   3 (bases 1 to 9658)

  AUTHORS   Tanabe,Y. and Sakamoto,N.
  TITLE   Direct Submission
  JOURNAL   Submitted (03-SEP-2003) Yoko Tanabe, Tokyo Medical and Dental
            University, Department of Gastroenterology and Hepatology; 1-5-45
            Yushima, Bunkyo-ku, Tokyo 113-8519, Japan
            (E-mail:ytanabe.gast@tmd.ac.jp, Tel:81-3-5803-5877,
            Fax:81-3-5803-0268)

FEATURES   Location/Qualifiers
     source   1..9658
                     /organism="Hepatitis C virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:11103"
     CDS   342..2840
                     /codon_start=1
                     /product="Fusion protein, Feo"
                     /protein_id="BAD00047.1"
                     /db_xref="GI:37991672"
                     /translation="MSTNPKPQRKTKGRAMEDAKNIKKGPAPPYPLEDGTAGEQLHKA
                     MKRYALVPGTIAFTDAHIEVDIYAAYPEMSVRLAEAMKRYGLNTNHRIKVCSENSLQ
                     FPMPVLGALFIGVAVAPANDIYNERELLNSMGISQPTVFVFGVSKKGLQKILNVQKKLPI
                     IQKIIMDSKDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTTALIMNSSGSTG
                     LPKGVALPHRTACVRRSHARDPIFGNQIIPDTAILSVVPFHHGFGMETTLGYLIICGFR
                     VVIMYRFEELFLRSLQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPL
                     SKEVGEAVAKRPHLPGYGLTETTSAILITPEGDDKPGAVGKVPFFEAKVVDLD
                     TGKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFIV
                     DRLKSLIKYKGYQVAPAELESTLLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTM
                     TEKEIVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKIAVR
                     RAMIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGA
                     INELQDEAAARLSWLATTGVPCAAVLDVTEAGRDWLLLGEVPGQDLLSSHLAPAEKVS
                     IMADAMRRLHTLDPATCPFDHQAKRIERARTRMEAGLVDQODLDEEHQGLAPAELFA
                     RLKARMPDGEDLVVTHGDACLPNIMVENGRRSGFIDCGRLGVADRYQDIALATRDIAE
                     ELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"

ORIGIN
Query Match   71.9%; Score 1992.2; DB 14; Length 9658;
Best Local Similarity   98.8%;
Matches 2022; Conservative 0; Mismatches 13; Indels 12; Gaps 1;

```
Qy    628  ACGGCGGATCTTTCCGCCCTTCTTCGGCCTTTATGAGGATCTCTCTTGATTTTTCTTGCGTCG  687
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2035  ACGGCGGATCTTTCCGCCCTTCTTCGGCCTTTATGAGGATCTCTCTTGATTTTTCTTGCGTCG  1976

Db   2628  AGTTTTCCGGTAAGAACCTTTCGGTTACTTCGTTCGTTCCACAAACACAACTCCTCCGCGCAACTTT  747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    688  AGTTTTCCGGTAAGAACCTTTCGGTTACTTCGTTCGTTCCACAAACACAACTCCTCCGCGCAACTTT  1916
Qy   1975  AGTTTTCCGGTAAGAACCTTTCGGTTACTTCGTTCGTTCCACAAACACAACTCCTCCGCGCAACTTT  1916

Qy    868  GCGCACACCTGCGTCGGACAGATGTTGGGGTGTTGGAGCAAGGATGGATTCAATTCAGCGGGA  927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1795  GCGCACACCTGCGTCGGACAGATGTTGGGGTGTTGGAGCAAGGATGGATTCAATTCAGCGGGA  1736

Qy   1855  CCGTGCTCCAAAACAACAACGGCGGCGGGAAGTTCACCCGGCGGTCATCGTCGGGGAAGACCT  1796
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808  CCGTGCTCCAAAACAACAACGGCGGCGGGAAGTTCACCCGGCGGTCATCGTCGGGGAAGACCT  867

Qy    748  TTCGCGGTTGTTACTTGACTGGGGACGTAATCCACGACGATCTCTTTTTCCGGTCATCGTCTTT  807
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1915  TTCGCGGTTGTTACTTGACTGGGGACGTAATCCACGACGATCTCTTTTTCCGGTCATCGTCTTT  1856

Qy    928  GCCACCTGATAGCCTTTGTACTTAATCAGAACTTCAGGCGGTCAAGCAGATGAAGAAGTGT  987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1795  GCCGCACCCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGTCAAGCGATGAAGAAGTGT  1736

Qy   1735  GCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGTCAAGCGATGAAGAAGTGT  1676

Qy    988  TCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCATCCTTGTCAATCAAG  1047
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1675  TCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCATCCTTGTCAATCAAG  1616

Qy   1048  GCGTTGGTGCGCTTCCGGATTGTTTTACATAACCGGACATAATCATTAGGAGCCTTCTCACACAC  1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1615  GCGTTGGTGCGCTTCCGGATTGTTTTACATAACCGGACATAATCATTAGGAGCCTTCTCAC  1556

Qy   1108  AGTTCGCCCTCTTTGATTAACGCCCAGCCCGGTTTCCGGTATCCAGATCCACACAACCTTGCT  1167
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1555  AGTTCGCCCTCTTTGATTAACGCCCAGCCCGGTTTCCGGTATCCAGATCCACACAACCTTGCT  1496
```

```
Qy  1168 TCAAAAAATGGAACAAACTTTACCGACCGCGCCCCGGTTTATCATCCCCCTGGGTGTAATC 1227
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1495 TCAAAAAATGGAACAAACTTTACCGACCGCGCCCCGGTTTATCATCCCCCTGGGTGTAATC 1436

Qy  1228 AGAATAGCTGATGTAGTCTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATGCGCGCCGGAC 1287
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1435 AGAATAGCTGATGTAGTCTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATGCGCGCCGGAC 1376

Qy  1288 CTCTTGGCAACCGCGTTCCCGGACTTCCTTAGAGAGGGGGAGCGCCACCAGAAGCAATTCG 1347
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1375 CTCTTGGCAACCGCGTTCCCGGACTTCCTTAGAGAGGGGGAGCGCCACCAGAAGCAATTCG 1316

Qy  1348 TGTAAATTAGATAAATCGTATTGTCAATCAGAGTGCTTTTGGGCCAAGAAGGAGAATAGG 1407
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315 TGTAAATTAGATAAATCGTATTGTCAATCAGAGTGCTTTTGGGCCAAGAAGGAGAATAGG 1256

Qy  1408 GTTTGGCACCAGCACCGCACTTTGAATCTTGTTAATCCTGAAGGCTCCCTCAGAAACAGCTCT 1467
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255 GTTTGGCACCAGCACCGCACTTTGAATCTTGTTAATCCTGAAGGCTCCCTCAGAAACAGCTCT 1196

Qy  1468 TCTTCAAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATACGTAGTAGTA 1527
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1195 TCTTCAAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATACGTAGTAGTA 1136

Qy  1528 AACATTCCAAAACCGTGATGGAAATGGAACAACACTTAAATATCCGGAATTGATT 1587
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1135 AACATTCCAAAACCGTGATGGAAATGGAACAACACTTAAATATCCGGAATTGATT 1076

Qy  1588 TGATTGCCAAAAATAGGATCTCTGGCATCCGAGAATCTCACGCAGGCAGTTCTATGAGGC 1647
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1075 TGATTGCCAAAAATAGGATCTCTGGCATCCGAGAATCTCACGCAGGCAGTTCTATGAGGC 1016

Qy  1648 AGAGCGACACCTTTAGGCAGACCAGTAGTCCAGGAGGTTCAATGATCAGTGCAATTGTC 1707
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1015 AGAGCGACACCTTTAGGCAGACCAGTAGTCCAGGAGGTTCAATGATCAGTGCAATTGTC 956

Qy  1708 TTGTCCCTATCGGAAGGACTCTGGCACAGAGTTCATTCATTAAAAACCGGGAGTAGATGA 1767
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  955  TTGTCCCTATCGGAAGGACTCTGGCACAGAGTTCATTCATTAAAAACCGGGAGTAGATGA 896

Qy  1768 GATGTGACGAACGTGTACATCGACTGAAAATCCCTGGTAATCCGTTTTAGAAATCCATGATA 1827
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  895  GATGTGACGAACGTGTACATCGACTGAAAATCCCTGGTAATCCGTTTTAGAAATCCATGATA 836

Qy  1828 ATAATTTTTTGGATGATGGGGAGCTTTTTTGCAGTTCAAAATTTTTTGCAACCCCTTT 1887
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  835  ATAATTTTTTGGATGATGGGGAGCTTTTTTGCAGTTCAAAATTTTTTGCAACCCCTTT 776

Qy  1888 TTGGGAAACGAACACCACGGTAGGCGTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCA 1947
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  775  TTGGGAAACGAACACCACGGTAGGCGTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCA 716

Qy  1948 TTATAAAATGTCGTTCGCGGCGGGGCGGTCCGATAAATAACGCGCCCACACCGGC 2007
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  715  TTATAAAATGTCGTTCGCGGCGGGGCGGTCCGATAAATAACGCGCCCACACCGGC 656

Qy  2008 ATAAAGAGAATTGAAGAGAGTTTTCACTGCATACGACGATTCTGTGATTTGTATTCAGCCCA 2067
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  655  ATAAAGAGAATTGAAGAGAGTTTTCACTGCATACGACGATTCTGTGATTTGTATTCAGCCCA 596

Qy  2068 TATGCGTTTCATAGCTTCTGCCAACCGAACGGGCACATTCGGAAGTACTCAGGGTAAGTAGTG 2127
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  595  TATGCGTTTCATAGCTTCTGCCAACCGAACGGGCACATTCGGAAGTACTCAGGGTAAGTAGTG 536

Qy  2128 TCCACCTCGATATGTGCACTCGTAAAAGCAATTGTTCCAGGAACCAGGCCGTATCTCTTC 2187
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  535  TCCACCTCGATATGTGCACTCGTAAAAGCAATTGTTCCAGGAACCAGGCCGTATCTCTTC 476

Qy  2188 ATAGCCTTATGTCAGTTGCTCTCCAGCGGTTCCATCTTCCAGCGGATAGAATGGGGCGCCGGG 2247
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  475  ATAGCCTTATGTCAGTTGCTCTCCAGCGGTTCCATCTTCCAGCGGATAGAATGGGGCGCCGGG 416

Qy  2248 CCTTTCTTTATGTTTTTGGGGTTCCATGGGCACGTCGGGTGGTTACGTTTGGTTTTT 2307
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  415  CCTTTCTTTATGTTTTTGGGGTTCCATGGGCACGTCGGGTGGTTACGTTTGGTTTTT 368

Qy  2308 CTTTGAGGGTTTAGGATTCGTTGCTCATGGATCGCAGGCGGTCTTACGAGGACTCCCCGGG 2367
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  367  CTTTGAGGGTTTAGGATTCGTTGCTCATGGATCGCAGGCGGTCTTACGAGGACTCCCCGGG 308

Qy  2368 CAAGCACCCTATCAGGCAGTACCACAAGGCCGTTCGGCAGACCCAACACTACTCGGGTAGCA 2427
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  307  CAAGCACCCTATCAGGCAGTACCACAAGGCCGTTCGGCAGACCCAACACTACTCGGCTAGCA 248

Qy  2428 GTCTTGCGGGGGGCACGCCCAAATCTTCCAGGGGTTATCCAAGGAAAGGACCC 2487
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  247  GTCTCGGCGGGGGGCACGCCCAAATCTTCCAGGGGTTATCCAAGGAAAGGACCC 188

Qy  2488 GGTCGTCCTGCAATTCCGGTGTACTCCACGGGGAGACCACTACTATGGCTCTCCCGGG 2547
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  187  GGTCGTTCCTGGCAATTCCGGTGTACTCCACGGGGAGACCACTACTATGGCTCTCCCGGG 128

Qy  2548 AGGGGGGGGTCCTGGAGGGCTGCACCACACTCATACTACTACCCATAGCTAGACGCTTTCTGC 2607
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127  AGGGGGGGGTCCTGGGAGGGCTGCACGACACTCATACTACTAACGGCCATGGCTTTCTGC 68

Qy  2608 GTGAAGACAGTAGTTCCTCACGAGGGGGAGTGAATTCATGGTGGAGTGTCGCCCCCATCAGGG 2667
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  67   GTGAAGACAGTAGTTCCTCACAGGGGAGTGATTCTATGGTGGAGTGTCGCCCCCAATCGGG 8

Qy  2668 GGCTGGC 2674
         |||||||
Db  7    GGCTGGC 1
```

RESULT 6

```
LOCUS       AB037684/c
ACCESSION   AB037684                5917 bp    DNA    circular SYN 21-MAY-2003
DEFINITION  Cloning vector pPLuc, complete sequence.
ACCESSION   AB037684
VERSION     AB037684.1  GI:7415874
KEYWORDS    PLuc; luciferase; modified C.
SOURCE      synthetic construct
ORGANISM    synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 5917)
AUTHORS     Chen,A.B., Kao,A.Y.-F. and Brown,C.M.
TITLE       A short open reading frame within the encapsidation signal affects
            the translation of the polymerase gene from hepatitis Bvirus
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 5917)
AUTHORS     Chen,A.B., Kao,A.Y.-F. and Brown,C.M.
TITLE       Direct Submission
JOURNAL     Submitted (27-JAN-2000) Chris M. Brown, University of Otago,
            Department of Biochemistry; P.O.Box 56, Dunedin, Otago 9001, New
            Zealand (E-mail:chris.brown@stonebow.otago.ac.nz,
            Tel:+64-3-479-7875, Fax:+64-3-479-7866)
FEATURES             Location/Qualifiers
     source          1..5917
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     misc_feature    1..495
                     /note="HBV Adw 5' leader of pregenomic RNA"
     CDS             35..94
                     /codon_start=1
                     /note="unnamed protein product; ORF"
                     /evidence=experimental
                     /transl_table=11
                     /protein_id="BAA93573.1"
                     /db_xref="GI:7415875"
                     /translation="MSHCSSLQAVPWVALGHGH"
     CDS             84..581
                     /codon_start=1
```

```
                    /transl_table=11
                    /product="modified C"
                    /protein_id="BAA93574.1"
                    /db_xref="GI:74158076"
                    /translation="MIIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALES
                    PEHCSPHTALRQAILCWGELMTLATWVGNNLEDPATRDLVVNTVNTMGLIRQLLW
                    FHISCLTFGRETVLEYLVSFGVWIRTPPAYRPNARRRRQKHERPEGAILSAGRWNR
                    WRATA"

     gene            490..2151
                    /gene="Pluc"

     CDS             490..2151
                    /gene="Pluc"
                    /note="P codon luciferase fusion"
                    /codon_start=1
                    /evidence=experimental
                    /transl_table=11
                    /product="luciferase"
                    /protein_id="BAA93575.1"
                    /db_xref="GI:74158077"
                    /translation="MPDVEDAXNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIA
                    FTDAHIEVDITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIG
                    VAVAPANDIYNERELLNSMGISOPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTD
                    YQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTA
                    CVRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELF
                    LRSLQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIAGGGAPLSKEVGEAVAKRF
                    HLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGE
                    LCVRGPMIMSGYVNNPEATNALIDKDGWLHSGDIAVWDEDEHFFIVDRLKSLIKYKGY
                    QVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQ
                    VTTAKKLRGGVFVDEVPKGLTGKLDARKIRETLIKAKKGGKIAV"

     promoter        5899..5917
                    /note="derived from pGLControl (Promega)"
                    /function="T7 promoter"

     ORIGIN

Query Match    Best Local Similarity    64.0%;   Score 1773.2;  DB 12;  Length 5917;
Matches 2019;  Conservative  86.4%;  Pred. No. 0;  Mismatches 258;  Indels  61;  Gaps  3;
```

Qy   1 GGATCCGCTGTGTGGAAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCACG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2835 GGATCCGCTGTGTGGAAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 2776

Qy   61 CAGAAGTATGCAAAGCAATGCATCTCAATTAGTCAGCAACACCAGTGTGTGGAAAGTCCCCAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2775 CAGAAGTATGCAAAGCAATGCATCTCAATTAGTCAGCAACACCAGTGTGTGGAAAGTCCCCAG 2716

Qy   121 CTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2715 CTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 2656

Qy   181 GCCCCTAACTCCGCCCATCCCGCCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2655 GCCCCTAACTCCGCCCATCCCGCCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA 2596

Qy   241 TGGCTGA------------CTAATTTTTTTTATTTATTGCAGAGG------------CCGAGGCC 280
            ||||||| |||||||||||||||||||||||||| |||||||
Db   2595 TCGTTCAGATCCTTATCGATTTTACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACC 2536

Qy   281 GCCTCGGCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTGTTTGTTTAACTTGT 340
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2535 GAGGGGAGGCGCCACCGAGGAGGAGGGGTGGGAGGAGAATAGGGTATGTAAATCTTGTA 1439

Qy   341 TGCAAAAGCTTACATGATCTGCAGAGATGAAAACATAAAATGAATGCAGCACTCTCTGCAG 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   341 ... 400

Qy   401 CTCACGGACCTTTCACACGACCTAGCCGTGACTAGGGCTAGATGGAGCCACCATTAAAGAAG 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1335 ... 1276

Qy   461 GAAGGAAAAGAAAGGAAAAGAGAGGAAAGGAAAAGAAAAA 495
            ||||||||||||||||||||||||||||||||||||||||
Db   1500 ... 495

```
Qy   1560  ACTTAAAATCGCAGTATCCGAATCATTTGATTGCAAAAATAGGATCTTGGCATGCGA  1619
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1215  ACTTAAAATCGCAGTATCCGGAATCATTTGATTGCAAAAATAGGATCTTGGCATGCGA  1156

Qy   1620  GAATCTCACGCAGCCAGTCTTCTATGAGGCAGAGGACGACACCTTTAGCACCAGTAGATCC  1679
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1155  GAATCTCACCCAGCCAGTTCTATGAGGCAGAGCGACACCTTTAGCACCAGTAGATCC  1096

Qy   1680  AGAGGAGTTCATGATCAGTAGTGCAATTGTCTCCCTATCGAAGGACTCTGGCACAAAATC  1739
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1095  AGAGGAGTTCATGATCAGTAGTGCAATTGTCTCCCTATCGAAGGACTCTGGCACAAAATC  1036

Qy   1740  GTATTCATTAAAACCGGAGGTAGATGAGATGAGAGATGACGAACGTGTACATCGACTGAAATCC  1799
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035  GTATTCATTAAAACCGGGAGGTAGATGAGAGATGACGAACGTGTACATCGACTGAAATCC  976

Qy   1800  CTGGTAATCCGTTTTATGAGATCCATGGATAATAAATATTTTTGGAGATCGATTGGAGAGCTTTTTTG  1859
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   975   CTGGTAATCCGTTTTTAGAGATCCATGGATAATAAATATTTTTGGAGATTGGGAGCTTTTTTTG  916

Qy   1860  CACGTTCAAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCAGGGTAGGCTGGGAAAT  1919
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   915   CACGTTCAAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCAGGGTAGGCTGGGAAAT  856

Qy   1920  GCCCATACTTGTTGACCAATTCACGTTCATTATAAAATGTCGTTCCGGGGGCCAACTGCAAC  1979
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   855   GCCCATACTTGTTGACCAATTCACGTTCATTATAAAATGTCGTTCCGGGGCCAACTGCAAC  796

Qy   1980  TCCGGATAAAATAAACCGCGCCCAAACACCGGCATAAGAAATCGAAGAGAGTTTCACTGCATAC  2039
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   795   TCCGGATAAAATAACCGCGCCCAACACCGGCATAAAGAAATTGAAGAGAGTTTCACTGCATAC  736

Qy   2040  GACGAATTCTGTGATTGTGTATTCAGCCCATAATCGTTTCAGTACGTTTCATAGCTTCTGCCAAACCGGA  2099
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   735   GACGAATTCTGTGATTGTGTATTCAGCCCATAATCGTTTCATAGCTTCTGCCAAACCGGA  676

Qy   2100  CATTTCGAAGTACTCAGCGTAAGTCATGTGTCCACCTCGATATGTGCATCGTGTAAAAGCAAT  2159
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   675   CATTTCGAAGTACTCAGCGTAAGTCATGTATGTCCACCTTCGATATGTGCATCGTGTAAAAGCAAT  616

Qy   2160  TGTTCCAGGAAACCAGGGCGTATCTCTTCATAGCCTTATGCGTTGCTCTCCCAGCGGTTCC  2219
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   615   TGTTCCAGGAAACCAGGGCGTATCTCTTCATAGCCTTATGCGTTGCTCTCCCAGCGGTTCC  556

Qy   2220  ATCTTCCAGCGGATAGAATGGCCCGGCGGATAGAATGGCCGCGGCCTTTCTTATGTTTTGGGCGTCTTCCATG  2277
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   555   ATCTTCCAGCGGATAGAATGGCCCGGCGGCCTTTCTTATGTTTTGGGCGTCTTCCATG  498
```

LOCUS   AF416989   6192 bp   DNA   circular SYN 15-APR-2002
DEFINITION Synthetic construct plasmid pTRE_hTF_Luc, complete sequence.
ACCESSION   AF416989
VERSION   AF416989.1 GI:17530181
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            other sequences; artificial sequences.

REFERENCE   1 (bases 1 to 6192)
  AUTHORS   Holen,T., Amarzguioui,M., Wiiger,M.T., Babaie,E. and Prydz,H.
  TITLE   Positional effects of short interfering RNAs targeting the human
            coagulation trigger Tissue Factor
  JOURNAL   Nucleic Acids Res. 30 (8), 1757-1766 (2002)
  MEDLINE   21935767
   PUBMED   11937629
REFERENCE   2 (bases 1 to 6192)
  AUTHORS   Holen,T., Amarzguioui,M. and Prydz,H.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-SEP-2001) The Biotechnology Centre of Oslo,
            University of Oslo, Gaustadalleen 21, Oslo 0373, Norway
FEATURES   Location/Qualifiers
     source   1..6192
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /plasmid="pTRE_hTF_Luc"
     misc_feature   1..318
                     /note="Tet-responsive element"
     CDS   481..3018
                     /codon_start=1
                     /transl_table=11
                     /product="tissue factor/luciferase fusion protein"
                     /protein_id="AAL40737.1"
                     /db_xref="GI:17530182"
                     /translation="METPAWPRVPRPETAVARTLLLGWVFAQVAGASGTNTVAAYNL
                     TWKSTNFKTILEWEPKPVNQVYTVQISTKSGDWKSKCFYTTDTECDLTDEIVKDVKQT
                     YLARVFSYPAGNVESTGSAGEPLYENSPEFTPYLETNLGQPTIQSFEQVGTKVNTVVE
                     DERTLVRRNNTFLSLRDVFGKDLIYTLYYWKSSSSGKKTAKTNTNEFLIDVDKGENYC
                     FSVQAVIPSRTVNRKSTDSPVECMGQEKGEFREIFYIIGAVAFVVIILVIILAISLHK
                     CRKAGVGQSWKENSPLGGPMEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGT
                     IAFTDAHIEVDITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALF
                     IGVAVAPANDIYNERELLNSMGISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSK
                     TDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHR
                     TACVRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEE
                     LFLRSLQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAK
                     RFHLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQR
                     GELCVRGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYK
                     GYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVA
                     SQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKIAV"
     misc_feature   481..1155
                     /note="derived from Homo sapiens tissue factor mRNA
                     sequence, GenBank Accession Number M16553"
     misc_feature   1366..3533
                     /note="derived from cloning vector pGL3-Enhancer sequence,
                     GenBank Accession Number U47297; contains luciferase gene"
     CDS   complement(5132..5992)
                     /note="confers ampicillin resistance"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /protein_id="AAL40738.1"
                     /db_xref="GI:17530183"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSONDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"

ORIGIN

Query Match   63.7%; Score 1766; DB 12; Length 6192;
Best Local Similarity   87.8%; Pred. No. 0;
Matches 2006; Conservative   0; Mismatches 170; Indels 108; Gaps   3;

```
Qy      1  GGATCCGGCTTGGGAATGCGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3533  GGATCCGCGTTGGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG  3474

Qy     61  CAGAAGTATGCAAAGCATCGCATCTCAATTAGTTCAGCAACCAGGTGTGGAAAGTCCCCAGG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3473  CAGAAGTATGCAAAGCATCGCATCTCAATTAGTTCAGCAACCAGGTGTGGAAAGTCCCCAGG  3414

Qy    121  CTCCCCAGCAGGCAGAAGTATGCAAAGCATGCAATCATTCAATTAGTTCAGCAACCAGTAGTTCC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3413  CTCCCCAGCAGGCAGGAAGTATGCAAAGCATGCAATCATTCAATTAGTTCAGCAACCAATAGTTCC  3354

Qy    181  GCCCCCTAACTCCGCCCCATCCCGCCCCCTAACTCACCCCCAGTTCCGCCCCATTCTCCGCCCCA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3353  GCCCCTAACTCCGCCCCATCCCGCCCCTAACTCCGCCCCAGTTCCGCCCCATTCTCCGCCTCCA  3294

Qy    241  TGGGTGACTAAATTTTTTTATTTATGCAGAGGCCGGAGGCCGCGGCCTCGAGCTATT  300
           |||||           |                    ||  ||  |||
Db   3293  TCG---------------------------------------TTCAGATCCTTATCGATT  3273

Qy    301  CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTACATGATC  360
```

```
REFERENCE     1
  AUTHORS    Yang,A.L. and Festing,M.
  TITLE      Methods and materials to determine the binding of p53 to a vector
             determining the binding of p53 to a vector
  JOURNAL    Patent: WO 0196602-A 3 20-DEC-2001;
             MEDICAL RESEARCH COUNCIL (GB)
FEATURES             Location/Qualifiers
  source        1..5064
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="pGL3Enhancer Vector Sequence"
ORIGIN

Query Match           63.6%;  Score 1763.2;  DB 6;  Length 5064;
Best Local Similarity  87.9%;
Matches  2002;  Conservative  0;  Mismatches  168;  Indels 108;  Gaps  3;
```

```
Qy      1 GGATCCGGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2255 GGATCCGGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 2196

Qy     61 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGGTGTGGGAAGTCCCAGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2195 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGGAAGTCCCAGG 2136

Qy    121 CTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2135 CTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 2076

Qy    181 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2075 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCTCCA 2016

Qy    241 TGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGCCGCCCTGGGGTCTCGACTATT 300
          ||||                          ||||||||||||
Db   2015 TCG-----------------------------TTCAGATCCTTATCGATT 1995

Qy    301 CCAGAAAGTGAGGAGGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTACATGATC 360
          ||| || ||  ||||
Db   1994 TTACCCACATTTGTAGAGGTTTTACTTG------------------------------ 1968

Qy    361 TGCAGAGAGGCCCAGTATCAGCACTCTCTGCGAGTCATGCGGCTCACGGGACCTTTCACAGCT 420
          ||||
Db   1967 ------------CTTTAAAAAACCTCCACACTCCCCCGAAGTGAAACATAAAATGAA 1919

Qy    421 AGCCCGTGACTAGGGCTAAGATGGAGCCACCATTAAAGAAGGAAGGAAAAGAAAGGAAAA 480
          ||||
Db   1918 TGCAATTGTTGTGTTAACTTGTTTATTGCAGCTTATATGGTTAACAAATAAAGCAAATAG 1859

Qy    481 AGAAGGAAAAGAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAGAAAAAAAAAAA 540
          |||
Db   1858 CATCACAAATTTCACAAATAAAGCACATTTTTTTCACTGCAATTCTTACTTGTGTTCCAA 1799

Qy    541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAATGGGCCTAAGA 600
          |||
Db   1798 ACTCATCAATGTATCTTATCA--------------------TGTCTGCTCGAAGC 1764

Qy    601 GGCCCGGAGTGTTTACCCCAACCTTTAAAGGGGGCCGGATCTTTCCGGCCTTTATG 660
          ||||||||||||||||||||||||||||||||||
Db   1763 GGCCGGCCCCCCGGACTCTAGAATTACACGGGACTTCTTTCCGGCCTTTATG 1704

Qy    661 AGGATCTCCTCTGATTTTTCTTTCCGGTAAGACCTTTCCGGTTACTTCGTCC 720
          ||||||||||||||||||||||||||||||||
Db   1703 AGGATCTCTCTGATTTTTTCTTCGGCCGTAGTTTTCCGGTTTTCCGTTAAGACCTTTCCGGTTTTCCGGTTTTCCGTTAAGACCTTTCCGTAGTTCGTCC 1644

Qy    721 ACAAAACACAACTCCTCCCCGGCAACTCTTTTACTTGACTGGCACGGTAATCC 780
          |||
Db   1643 ACAAACACAACTCCTCCCCGGGCAACTCTTTTCGCGGTTTTACTTGACTGGCACGGTAATCC 1584

Qy    781 ACGGATCTCTTTTTTCCGTCATCGTCTTTCCGGTGTCTCCAAAAACAACAACGGCGGCGGGGAAGT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1583 ACGGATCTCTTTTTTCCGTCATCGTCTTTCCGGTGTCTCCAAAAACAACAACGGCGGCGGGGAAGT 1524
```

```
Qy    841 TCACCCGGCGTCATCGTCGGGGAAGACCTGCGACACCTGCCGGTCGGGAAGACCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1523 TCACCCGGCGTCATCGTCGGGGAAGACCTGCGACACCTGCCGGTCGGGAAGACCG 1464

Qy    901 AGCAAGATGGAATTCCAAATTCAGCGCGGAGCCCACCCTGATAGCCTTTGTACTTAATCAGAGAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1463 AGCAAGATGGAATTCCAAATTCAGCGCGGAGCCCACCCTGATAGCCTTTGTACTTAATCAGAGAC 1404

Qy    961 TTCAGGCGGGTCAAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTTGCTTGCTTCCCAGAA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1403 TTCAGGCGGGTCAAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAA 1344

Qy   1021 TGTAGCCATCCATCCTTGTCAAATCAAGGACGGTTGGTCGCTTCCGGCTTCCGGGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1343 TGTAGCCGCCATCCATCCTTGTCAAATCAAGGGCGGTTGGTCGCTTCCGGGATTGTTTTACATAACCG 1284

Qy   1081 GACATAAATCATTAGGACCTCTCCACACACAGTTCGCCTCTTGATTAACGCCCAGCCGTTTTC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1283 GACATAAATCATTAGGACCTCTCCACACACAGTTCGCCTCTTGATTAACGCCCAGCCGTTTTC 1224

Qy   1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTACCGACGGGCCC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTACCGACGGGCCC 1164

Qy   1201 GGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGTTGCTCAGTGAGTCCCATAT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 GGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGTTGCTCAGTGAGTCCCATAT 1104

Qy   1261 CCTTCGCCTGATACCTTGGGAGTTGGCAGATGGAAACCTCTTCGGAACCCTCCCGGTTCCTTAGAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 CCTTCGCCTGATACCTTGGGAGTTGGCAGATGGAAACCTCTTCGGAACCCTCCCGGTTCCTTAGAG 1044

Qy   1321 AGGGGAGGGCGCCACCAGGAAGCAAATTCGTGTAAAATTAGATAAAATCGTTCAATCAGA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 AGGGGAGCGCGCCACCAGGAAGCAAATTCGTGTAAAATTAGATAAAATCGTTCAATCAGA 984

Qy   1381 GTGCTTTGGCGAAGGAAGAGAGATAGGGTTGCACCGACCGCACGCTTTGAATCTTGTAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    983 GTGCTTTGGCGAAGGAAGAGAGATAGGGTTGGCACCGACCGCACTTTGAATCTTGTAA 924

Qy   1441 TCCCTGAAAGGCTCCCCAGATCCGAGTATCCGAAATCAGCTTCTCTTTCAAAATCTATACATTAAGTACGAGACTGGAAAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    923 TCCCTGAAAGGCTCCCCAGATCCGAGTATCCGAAATCAGCTTCTCTTTCAAAATCTATACATTAAGACGAGACTCGAAAT 864

Qy   1501 CCACATATCAAATATCCGAGTGTAGTAACAATTCCAAAACCGTCGATGATGGAACAACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    863 CCACATATCAAATATCCGAGTGTAGTAACAATTCCAAAACCGTCGATGATGGAACAACA 804

Qy   1561 CTTAAAATCGCAGTATCCGGAATGTGATTGATTGAGATCTCTGGCATCCGAG 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    803 CTTAAAATCGCAGTATCCGGAATGTGATTGATTGAGATCTCTGGCATCCGAG 744

Qy   1621 AAATCTTCACGCCAGGCGAGTTTCTATTGAGGGCAGCAGCGACCAGTTAGGATCCA 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 AAATCTTCACGCAGGCGAGTTTCTATTGAGGGCAGCAGCGACCAGTTAGGATCCA 684

Qy   1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCAAAATCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCG 624

Qy   1741 TATTCATTAAAACCCGGCGTAGACGTGACGAAGGCTGTACATCGACTGAAATCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 TATTCATTAAAACCCGGCGTGAGTGACGTGACATCGACTGAAATCCC 564

Qy   1801 TGGTAATCCGTTTTAGAATCCATGATATAAATATTTTTGAATGAATTGAAGCTTTTTTGC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATGATTTTTTGC 504

Qy   1861 ACGTTCAAAATTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATG 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 ACGTTCAAAATTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATG 444
```

```
RESULT 9

CVU47297/c

LOCUS       CVU47297                5064 bp    DNA     linear   SYN 17-APR-2002
DEFINITION  Cloning vector pGL3-Enhancer, complete sequence.
ACCESSION   U47297
VERSION     U47297.2  GI:13195705
KEYWORDS    .
SOURCE      Cloning vector pGL3-Enhancer
  ORGANISM  Cloning vector pGL3-Enhancer
            other sequences; artificial sequences; vectors.
REFERENCE   1  (bases 1 to 5064)
  AUTHORS   Groskreutz,D.J. and Schenborn,E.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1996) D.J. Groskreutz, R&D, Promega Corporation,
            5445 East Cheryl Parkway, Madison, WI 53711, USA
REFERENCE   2  (bases 1 to 5064)
  AUTHORS   Kenefick,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAR-2001) Technical Writing, Promega Corporation,
            2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARK      Sequence update by submitter
COMMENT     On Mar 5, 2001 this sequence version replaced gi:1200465.
FEATURES             Location/Qualifiers
     source          1..5064
                     /organism="Cloning vector pGL3-Enhancer"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:45859"
                     /note="luciferase reporter vector"
     misc_feature    1..58
                     /note="multiple cloning site"
     gene            88..1740
                     /gene="luc+"
     CDS             88..1740
                     /gene="luc+"
                     /codon_start=1
                     /product="luciferase"
                     /protein_id="AAB9086.1"
                     /db_xref="GI:1200466"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVDITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMGISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESLLLQHPNIFDAGVAGLPDDDAGELPAAVVLEHGKTMTKEIVDVVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKIAV"
ORIGIN

Query Match           63.6%;  Score 1763.2;  DB 12;  Length 5064;
Best Local Similarity   87.9%;
Matches 2002;  Conservative   0;  Mismatches 168;  Indels 108;  Gaps  3;
```

(sequence alignment data — Qy/Db paired rows)

RESULT 10
LOCUS   AX339208/c
                AX339208
DEFINITION   Sequence 2 from Patent WO0196602.
ACCESSION   AX339208
VERSION   AX339208.1   GI:18135469
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
                other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Yang,A.L. and Festing,M.
  TITLE   Methods and materials to determine the p53 status of a sample by
                determining the binding of p53 to a vector
  JOURNAL   Patent: WO 0196602-A 2 20-DEC-2001;
                MEDICAL RESEARCH COUNCIL (GB)
FEATURES   Location/Qualifiers
     source   1..5256
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pGL3-Control Vector Sequence"
ORIGIN

Query Match   63.6%; Score 1763.2; DB 6; Length 5256;
Best Local Similarity   87.9%;
Matches 2002; Conservative 0; Mismatches 168; Indels 108; Gaps 3;

(This page consists of a dense DNA sequence alignment between two sequences labeled "Qy" and "Db", arranged in paired rows with alignment bars and base-pair position numbers along both margins. The nucleotide sequences are printed as rotated, tightly-spaced text and are not legibly resolvable for faithful character-by-character transcription.)

```
Qy    2221 TCTTCCAGCCGGATAGAATGGCGCGCGCGCGGGCCCTTTTCTTTATGTTTGTTTTTTGGCGTCTTCCATGG 2278
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     335 TCTTCCAGCCGGATAGAATGGCGCGCCGCCCGGGCCCTTTCTTTCTTTATGTTTGTTTTTTGGCGTCTTCCATGG 278
```

```
Query Match            63.6%; Score 1763.2;  DB 12;  Length 5256;
Best Local Similarity  87.9%;  Pred. No. 0;  Mismatches 168;  Indels 108;  Gaps 3;
Matches 2002;  Conservative    0;  Mismatches 168;  Indels 108;  Gaps 3;

Qy       1 GGATCCGCGTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2447 GGATCCGCGTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 2388

Qy      61 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2387 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 2328

Qy     121 CTCCCCAGCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2327 CTCCCCAGCAGCAGGCAGAAGTATGCAAAGCATCTCAATTAGTCAGCAACCATAGTCCC 2268

Qy     181 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCGCCCCAGTTCCGCCCATTCTCCGCCCCA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2267 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCGCCCCAGTTCCGCCCATTCTCCGCCCCA 2208

Qy     241 TGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATT 300
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    2207 TCG------------------------------TTCAGATCTTTATCGATT 2187

Qy     301 CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTACATGATC 360
              ||| |||||||||||||||||||||||||||||||||||||||||||||
Db    2186 TTACCACATTGTAGAGGTTTTACTTG------------------------------ 2160

Qy     361 TGCAGAGAGGCCCAGTATCAGCGACTCTCTGCAGTCATGGCGCTCACGGACTTTCAGCT 420
           ||||||||||||||||||||||     ||||||||||||||||||||||||
Db    2159 ------------CTTTAAAAAACCTCCACAACCTCCCCCTGAACCTGAACACATAAAATGAA 2111

Qy     421 AGCCGTGACTAGGGCTAGCGATGGACGCCACCATTAAAGAAGAGGAGGAAAGAAGAAGAA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2110 TGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAG 2051

Qy     481 AGAAGGAAAGAGAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAGAAAAAAAAAA 540
           ||||              ||||||||||||||||||||||
Db    2050 CATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAA 1991

Qy     541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGGAAATGGGCTTAAGA 600
           |||||                                     |||||||
Db    1990 ACTTCATCAAATGTATCTTATCA---------------TGTCTGCTCGAAGC 1956

Qy     601 GGCCCGGAGTGTTACCCGAACCTTTAAACGGGCGATCTTTCCGCCCTTCTTCTGGCCTTTATG 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1955 GGCCGGCCCGCGCCGCGACTCTAGAATTACACGGCGATCTTTCCGCCCTTCTTGGCCTTTATG 1896

Qy     661 AGGATCTCTCTGATTTTTCTTGCGTCGCGAGTTTTCCGGTGAAGACCTTTCGGTTCGTCC 720
           ||||     ||| |||||||||| |||||||||||||||||||||||||||||||
Db    1895 AGGATCTCTCTGATTTTTCTTGCGTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCC 1836

Qy     721 ACAAACACAACTCCCTCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1835 ACAAACACAACTCCTCCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCC 1776

Qy     781 ACGATCTCTTTTTCCGTCATCGTCTTTCCGGTGCTCCAAAAACACACAACGGGGAAGT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1775 ACGATCTCTTTTTCCGTCATCGTCTTTCCGGTGCTCCAAAAACAACAACGGGGAAGT 1716
```

```
Qy    841  TCACCGGCGTCATCGTCGTCGGGACGAAGACCTGGTCGTCGAAGATGTTGGGGTGTTGG   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1715  TCACCGGCGTCATCGTCGTCGGGACGAAGACCTGGTCGTCGAAGATGTTGGGGTGTTGG  1656

Qy    901  AGCAAGATGGATTCCAATTCAGCGGGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGAC   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1655  AGCAAGATGGATTCCAATTCAGCGGGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGAC  1596

Qy    961  TTCAGGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTTCGTTCCAGA   1020
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1595  TTCAGGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTTCGTTCCAGA  1536

Qy   1021  TGTAGCCATCCATCCTTGTCAATCAAAGCGTGTGTTTTACATAACCG   1080
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1535  TGTAGCCATCCATCCTTGTCAATCAAAGCGTGTGTTTTACATAACCG  1476

Qy   1081  GACATAATCATAGGACCTTCTCACACACAGTTCGGCCTCTTTGATTAACGGCCCAGCGGTTTTC   1140
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1475  GACATAATCATAGGACCTTCTCACACACAGTTCGGCCTCTTTGATTAACGGCCCAGCGGTTTTC  1416

Qy   1141  CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAAATGGAACAACTTTACCGACCGGCCC   1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415  CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAAATGGAACAACTTTACCGACCGGCCC   278
```

(alignment continues — query/database nucleotide sequence rows spanning positions to approx. Qy 1921 / Db 636)

```
LOCUS       AF416988                6612 bp    DNA     circular SYN 12-DEC-2001
DEFINITION  Synthetic construct plasmid pTRE_PSKH1_Luc, complete sequence.
ACCESSION   AF416988
VERSION     AF416988.1  GI:17530178
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 6612)
  AUTHORS   Amarzguioui,M., Brede,G., Babaie,E., Grotli,M., Sproat,B. and
            Prydz,H.
  TITLE     Secondary structure prediction and in vitro accessibility of mRNA
            as tools in the selection of target sites for ribozymes
  JOURNAL   Nucleic Acids Res. 28 (21), 4113-4124 (2000)
   MEDLINE   20512557
   PUBMED    11058107
REFERENCE   2  (bases 1 to 6612)
  AUTHORS   Amarzguioui,M., Holen,T. and Prydz,H.P.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2001) The Biotechnology Centre of Oslo,
            University of Oslo, Gaustadalleen 21, Oslo 0373, Norway
FEATURES             Location/Qualifiers
     source          1..6612
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /plasmid="pTRE_PSKH1_Luc"
     misc_feature    1..318
                     /note="Tet-responsive element"
     CDS             511..3438
                     /codon_start=1
                     /transl_table=11
                     /product="protein serine kinase/luciferase fusion protein"
                     /protein_id="AAL40735.1"
                     /db_xref="GI:17530179"
                     /translation="MGCGTSKVLPEPKDVQLDLVKKVEPSGTKSDVYKHFITEVDS
                     VGPVKAGFPAASQYAHPCPGPTAGHTEPPSEPPRRARVAKYRAKEPDRVTAKYDIKA
                     LIGRGSFSRVRVEHRATRQYAIKMIETKVREGREVECESELRVLRRVRHANIIQLVE
                     VFETQERVYMVMELATGGELFDRIIAKGSFTERDATRVLQMVLDGVRYLHALGITHRD
                     LKPENLLYYHPCTDSKIIITDFGLASARKGDDCLMKTTCGTPEYIAPEVLVRKPYTN
                     SVDMWALGVIAYILLSGTMPFEDDNRTRLYRQILRGKYSYGGEPWPSVSNLAKDFDR
                     LLTVDPGARMTLQALRHPWVSMAASSMKNLHRSISQNLLKRASSRCQSTKSAQST
                     RSSRSTRSNKSRVRERELRELNLRYQQYNGAMEDAKNIKGPAPYFPLEDGTAGEQ
                     LHKAMRYALYLPGTIAFTDAHIEVDITYAEYFPEMSVRLAEAMKRYGLNTNHRIVVCSE
                     NSLQFFMPVLGALFIGVAVAPANDIYNERELLNSMGISQPTVFVFVSKKGLQKIINVQK
                     KLPIIQKIIIMDSKTDVQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSS
                     GSTGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMPTTLGYLI
```

```
                CGPRVVLMYRFEEELFLRSLQDYKIQSALLVPFLFSFPPAKSTLIDKYLDLSNLHEIASG
                GAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKV
                VDLDTGKTLGLVNORGELCVRGPMIMSGYVVNNPEATNALIIDKDGWLHSGDIAYWDEDEH
                FFIVDRLKSLIKYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVLEH
                GKTMTEKEIVDTVASQVTTAKKLRGGVVFVDEVPRGLTGKLDARKIREILIKAKKGGK
                IAV"

misc_feature    511..1782
                /note="derived from Homo sapiens protein serine kinase
                mRNA sequence, GenBank Accession Number Aj272212"

misc_feature    1784..3953
                /note="derived from cloning vector pGl3-Enhancer sequence,
                GenBank Accession Number U47297; contains luciferase gene"

CDS             complement(5552..6412)
                /note="AmpR; confers ampicillin resistance"
                /codon_start=1
                /transl_table=11
                /product="beta-lactamase"
                /protein_id="AAL40736.1"
                /db_xref="GI:17530180"
                /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                IELDLNGGKILSSFRPEERFPMMSTFKVLLCGAVLSRIDAGQEOLGRRIHYSQNDLVE
                YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGGPKELTAFLHNMGDHVIRL
                DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                EIGASLIKHW"

ORIGIN
```

```
Query Match          63.6%;     Score 1763.2;   DB 12;   Length 6612;
Best Local Similarity 87.9%;    Pred. No. 0;
Matches 2002; Conservative 0; Mismatches 168; Indels 108; Gaps 3;
```

```
Qy      1 GGATCGGCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3953 GGATCGGCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 3894

Qy     61 CAGAAGTATGCAAAAGCACATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3893 CAGAAGTATGCAAAAGCACATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 3834

Qy    121 CTCCCCAGCAGGCAGGCGAGGAGGTATGCAAAGCATCATCTCAATTAGTCAGCAACCATAGTCCC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3833 CTCCCCAGCAGGCAGGCGAGGAGGTATGCAAAGCATCATCTCAATTAGTCAGCAACCATAGTCCC 3774

Qy    181 GCCCCTCAACTCCCGCCCCATCCCGCCCCCAGTTCCGCCCCATTCTCCGCCCCA  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3773 GCCCCTAACTCCCGCCCCATCCCGCCCCCAGTTCCGCCCCATTCTCCGCCCCA 3714

Qy    241 TGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCGCCTCGCGGCTATT  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3713 TCG--------------------TTCAGATCCTTATCCGATT 3693

Qy    301 CCAGAAGTAGTGAGGAGGCTTTTTGGAGGCCTAGGCTTTTTGCAAAAAAGCTTACATGATC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3692 TTACCACATTTGTTAGAGGTTTTTACTTG------ 3666

Qy    361 TGCAGAGAGGCCAGTATCAGCACTCCTTCTGCAGTCATGCGGCTCAGGGACCTTTCACAGCT  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3665 -------CTTTAAAAAACTTCCCACACTTCCCCCTGAACCTGAACATAAAATGAA 3617

Qy    421 AGCGGTGACTAGGGCTTAGACATGATGAGCGCACCATTAAAAGAAGGAGGAAGAGGAAGAAA  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3616 TGCAATTGTTGTTGTTGTTAACTTGTTTATTATGATGGTTACAAATAAAAGCAATAG 3557

Qy    481 AGAAGGAGAAAGAAAAGAGAAAAAAAAAAGAAAAAAAAAAGAAAAAAAA  540
          |||||||||||||||||||||||||||
Db   3556 CATCACAAATTTCACAAATAAAGCAATTTTTTCACTGCATTCGTAGTTGTTGGGTTCCAA 3497

Qy    541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAAAATGCCCTAGA  600
          ||         |||||||||||||||||||||||||||||||||||
Db   3556 AATCTCACGGCAGGCAGAGTTCTATGAGGCGCAGAGCGACACCTTTAGGCAGAGATCCA 3462

Qy    601 GGCCGGGAGTGTTTACCCAAACCTTTAAACGGGCGATCTTTCCGCCCTTCTTGGCCTTTATG  660
          |||||
Db   3496 ACTCATCAATGTATCTTATCA------ 3462
```

```
Qy   1021 TGTAGGCCATCCATCTCCTTGTCAATCAAGCCGCTTCCGCGGATTGTTTTGTTTACCATAACCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 TGTAGGCCATCCATCTCCTTGTCAATCAAGCCGCTTCCGCGGATTGTTTTGTTTACCATAACCG 2982

Qy   1081 GACATAATCATTAGGACACTTCTCACACACACAGTTCGCCTCTTGTGATTAATGCACGTTTTC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 GACATAATCATTAGGACACTTCTCACACACACAGTTCGCCTCTTGTGATTAATCGACGTTTTC 2922

Qy   1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAAATGGAACAACTTTACCGGCCGCGCC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAAATGGAACAACTTTACCGGCCGCGCC 2862

Qy   1201 GGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTATCGTCTCAGTGAGCCATAT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 GGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTATCGTCTCAGTGAGCCATAT 2802

Qy   1261 CCTTGCCTGATACCTGGCAGATGGAACCTCTGGCAACCGCTTCCCGCACTTCCTTAGAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 CCTTGCCTGATACCTGGCAGATGGAACCTCTGGCAACCGCTTCCCGCACTTCCTTAGAG 2742

Qy   1321 AGGGGAGGCGCCCACCAGAAGCACAATTCGTTGTAAATTAGATAAAATCGTATTTCAATCAGA 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 AGGGGAGGCGCCCACCAGAAGCACAATTCGTTGTAAATTAGATAAAATCGTATTTCAATCAGA 2682

Qy   1381 GTGCTTTTGGCGAAGAAGAAATAAGGGTTGGCACCAGCGCACTTTGAAATCTTGTAA 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2681 GTGCTTTTGGCGAAGAAGAAATAAGGGTTGGCACCAGCGCACTTTGAAATCTTGTAA 2622

Qy   1441 TCCTGAAAGGCTTCCTCAGGAAAACAGCTCCTTCTTCAAATCTATACATTAAGAGGACTGAAAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2621 TCCTGAAAGGCTTCCTCAGGAAAACAGCTCCTTCTTCAAATCTATACATTAAGAGGACTGAAAT 2562

Qy   1501 CCACATATCAAAATATCCGACTGTAGTAGAAACATTCCAAAAACCGTGATGAGGAATGGAACACA 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2561 CCACATATCAAAATATCCGACTGTAGTAGAAACATTCCAAAAACCGTGATGAGGAATGGAACAACA 2502

Qy   1561 CTTAAAAATGCAGTATCCGGAATGATGTTTGCCAAAAAAATAGGATCTCTGGCATGCGAG 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2501 CTTAAAAATGCAGTATCCGGAATGATGTTTGCCAAAAAAATAGATCTCTGGCATGCGAG 2442

Qy   1621 AATCTCACGGACGGAGGCAGAGTTCTATGAGGCGCAGAGCGCAGAGTAGAATCCA 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2441 AATCTCACGGACGGAGGCAGAGTTCTATGAGGCGCAGAGCGCAGAGTAGAATCCA 2382

Qy   1681 GAGGAGTTCATGACTAGTGCAAATTGTCTTGTGTCCCTATCGAAGGACTCTCGGCACAAAATCG 1740
          ||||||
Db   2381 GAGGAGTTCATGACTAGTGCAAATTGTCTTGTGTCCCTATCGAAGGACTCTCGGCACAAAATCG 2322
```

```
RESULT 13

CQ803129/c

LOCUS       CQ803129        7014 bp    DNA    linear    PAT 10-MAY-2004
DEFINITION  Sequence 15 from Patent WO2004035782.
ACCESSION   CQ803129
VERSION     CQ803129.1  GI:47110086
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Seibler,J.; Schwenk,F.; Kuehn,R. and Kueter-Luks,B.
  TITLE     Sirna mediated gene silencing in transgenic animals
  JOURNAL   Patent: WO 2004035782-A 15 29-APR-2004;
            ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..7014
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: GAGGS-Fluc"
ORIGIN

Query Match     63.6%;  Score 1763.2;  DB 6;  Length 7014;
Best Local Similarity   87.9%;
Matches 2002;  Conservative    0;  Mismatches  168;  Indels  108;  Gaps   3;
```

```
Qy        1 GGATCCGCTGTGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCGACAGG    60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4183 GGATCCGCTGTGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCGAGG  4124
```

[Sequence alignment continues — Qy/Db paired rows throughout]

RESULT 14

```
LOCUS       AX528380               7788 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 11 from Patent WO0220749.
ACCESSION   AX528380
VERSION     AX528380.1  GI:25172613
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Owman,C.S., Olde,B.A. and Kotarsky,K.
  TITLE     Reporter system for cell surface receptor-ligand binding
  JOURNAL   Patent: WO 0220749-A 11 14-MAR-2002;
            Owman, Christer S. D. (SE) ; Olde, Bjorn A. (SE) ; Kotarsky, Knut
            (SE)
FEATURES             Location/Qualifiers
     source          1..7788
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="SEQUENCE OF pcFUSII PLASMID"
ORIGIN
```

Query Match   63.5%; Score 1758.2; DB 6; Length 7788;
Best Local Similarity   87.9%; Pred. No. 0; Mismatches 168; Indels 108; Gaps 3;
Matches 1997; Conservative   0; Mismatches 168; Indels 108; Gaps 3;

```

```
Db   565 CTCTCTGATTTTTCTTGCGTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCCACAAA 624

Qy   726 CACAACTCCTCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCCACGAT 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625 CACAACTCCTCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCCACGAT 684

Qy   786 CTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACAACAACGGCGGCGGAAGTTCACC 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685 CTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACAACAACGGCGGCGGAAGTTCACC 744

Qy   846 GGCGTCATCGTCGGGAAGACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGGAGCAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   745 GGCGTCATCGTCGGGAAGACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGGAGCAA 804

Qy   906 GATGGATTCCAATTCAGCGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAG 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805 GATGGATTCCAATTCAGCGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAG 864

Qy   966 GCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAATGTAG 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 GCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAATGTAG 924

Qy   1026 CCATCCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCGGACAT 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   925 CCATCCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCGGACAT 984

Qy   1086 AATCATAGGACCTCTCACACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTCCCGGT 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   985 AATCATAGGACCTCTCACACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTCCCGGT 1044

Qy   1146 ATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCCGGTTT 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 ATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCCGGTTT 1104

Qy   1206 ATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATATCCTTG 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 ATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATATCCTTG 1164

Qy   1266 CCTGATACCTGGCAGATGGAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAGAGGGG 1325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 CCTGATACCTGGCAGATGGAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAGAGGGG 1224

Qy   1326 AGCGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCT 1385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 AGCGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCT 1284

Qy   1386 TTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAATCCTG 1445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 TTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAATCCTG 1344

Qy   1446 AAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACA 1505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 AAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACA 1404

Qy   1506 TATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAA 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 TATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAA 1464

Qy   1566 AATCGCAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAGAATCT 1625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465 AATCGCAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAGAATCT 1524

Qy   1626 CACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCAGAGGA 1685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525 CACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCAGAGGA 1584

Qy   1686 GTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCGTATTC 1745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCGTATTC 1644

Qy   1746 ATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCCTGGTA 1805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1645 ATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCCTGGTA 1704

Qy   1806 ATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGCACGTT 1865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1705 ATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGCACGTT 1764

Qy   1866 CAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATGCCCAT 1925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1765 CAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATGCCCAT 1824

Qy   1926 ACTGTTGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACTCCGAT 1985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1825 ACTGTTGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACTCCGAT 1884

Qy   1986 AAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACGACGAT 2045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1885 AAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACGACGAT 1944

Qy   2046 TCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGACATTTC 2105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1945 TCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGACATTTC 2004

Qy   2106 GAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCC 2165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2005 GAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCC 2064

Qy   2166 AGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCATCTTC 2225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2065 AGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCATCTTC 2124

Qy   2226 CAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGG 2278
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2125 CAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGG 2177

RESULT 15
AX528381
LOCUS   AX528381   7312 bp   DNA   linear   PAT 21-NOV-2002
DEFINITION Sequence 12 from Patent WO0220749.
ACCESSION   AX528381
VERSION   AX528381.1 GI:25172614
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Owman,C.S., Olde,B.A. and Kotarsky,K.
  TITLE   Reporter system for cell surface receptor-ligand binding
  JOURNAL   Patent: WO 0220749-A 12 14-MAR-2002;
            Owman, Christer S. D. (SE) ; Olde, Bjorn A. (SE) ; Kotarsky, Knut
            (SE)
FEATURES   Location/Qualifiers
     source   1. .7312
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="SEQUENCE OF pcFUSII-IE PLASMID"
ORIGIN

  Query Match   63.1%; Score 1747.2; DB 6; Length 7312;
  Best Local Similarity   87.8%; Pred. No. 0;
  Matches 1997; Conservative   0; Mismatches 168; Indels 109; Gaps   4;

Qy   6 CGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13 CGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA 72

Qy   66 GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   73 GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC 132

Qy   126 CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133 CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCC 192
```

Qy    186  TAACTCCGCCCATCCGCGCCCCTAACTCCGCCCATTCTCCCGCCATGGCT   245
           |||||||||||||||||||||| ||||||||||||||||| ||||||||
Db    193  TAACTCCGCCCATCCGCGCCCTAACTCCGCCCAGTTCCCGCCCATTCCGGCTCG--   250

Qy    246  GACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCCTCGAGCTATTCCAGA   305
           |||||||||||||||||||||||||||||||||||||||||||  ||||||
Db    251  ------------------------------------TTCAGATCCTTATCGATTTACC   273

Qy    306  AGTAGTGAGGAGGGCTTTTTTGGAGGCCTTAGGCTTTTGCAAAAAGCTTACATGCTGCAG   365
           |||||   |||  |||||||||||||||||||||||||||||||||||
Db    274  ACATTTGTAGAGGTTTTACTTG--------------------------------   295

Qy    366  AGAGGCCAGTATCAGCACTTCTCTGCAGTCATGCGGCTCACGGACCTTTCACAGCTTAGCCG   425
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    296  -----CTTTAAAAAAACCTCCCACACCTCCCCTGAACCTGAACATAAAATGAATGCAA   349

Qy    426  TGACTAGGGCTAAGATGGGAGCCACCATTTAAAGAGGAAGGAAGAGAGAGAGAGAGAGAGAG   485
           |||||||||||||||||||||||||
Db    350  TTGTTGTTGTTTGTTAACTTGTTTTATTCGAGCTTTATAAATGGTTTACAAATAGCAATCA   409

Qy    486  GAAAGAGAAAAAAGGGAAAAAAAAAGAGAAAAAAGAAGATTTTTTTCACTGCATTCTAGTTGTTGT   545
           |||||||||||||||||||||||||||||||
Db    410  CAAAATTTCACAAATAAAGACATTTTTTTCACTGCATTCTAGTTGTTGTTCCAAAACTCA   469

Qy    546  AAAAAAAAAAAACAGGAAAATGGCCTAAGACCCCC   605
           |||||||||||||||
Db    470  TCAATGTAGTTACTTACA-----------TGTCTGCTCGAACGCGCGGCCCG   504

Qy    606  GAGTGTTTACCCAACCTTTTAAAGGGGATCTTTCCCGCCCCTTCTTGGCCTTTATGAGGAT   665
           |||||||||||||||||||||||||
Db    505  GCCGCCCCCGGACTCTTAGGAATTACACGGGGCGGATCTTTCCCGCCCCTTCTTGGCCTTTATGAGGAT   564

Qy    666  CTCTCTGATTTTTCTTTGCCGTCGAGTTTTTCCGGTAAGACCTTTCGGTTACTTCGTTCACAAA   725
           |||||||||||||||||||||||||
Db    565  CTCTCTGATTTTTCTTTGCCGTCGAGTTTTTCCGGTAAGACCTTTCGGTTACTTCGTTCACAAA   624

Qy    726  CACAACTCCTCCCGCCAACTTTTTTCGCGGTTGTTACTTGCTGGCGACGTAATCCACGAT   785
           |||||||||||||||||||||
Db    625  CACAACTCCTCCCGCCAACTTTTTTCGCGGTTGTTACTTGCTGGCGACGTAATCCACGAT   684

Qy    786  CTCCTTTTTCCGTCATCGTCTTTCCGGTCCGGTTCCAAAAACAACAACAAACCACC   845
           |||||||||||||||||||||
Db    685  CTCTTTTTCCGTCATCGTCTTTCCGGTGTCTCCAAAAACAACAACAAACCACC   744

Qy    846  GGCCTCATGTCGGGAAGAGACCTGCCACACCTGCCTCGAAGATGTTGGTGTTGGAGCAA   905
           |||||||||||||||||||||
Db    745  GGGCGTCATGTCGGGAAGAACCTGGGAAGACCTGGGAACACCTGCCTCGAAGATGTTGGGGTGTTGGAGCAA   804

Qy    906  GATGGAATTCCAATTCAGCGGGGAGGCCACCATGGCACTTGATTAGGCCTTTGTACTTAATCAGACTTCAG   965
           |||||||||||||||||||||
Db    805  GATGGATTCCAATTCAGCGGGAGCCCACCTGATAGCCTTTGTACTTAATCAGACTTCAG   864

Qy    966  GCGGTCAACGGATGAAGAAGTGTTGTGTTCGTCTTCGTCCCAGGAATGTAG   1025
           |||||||||||||||||||||
Db    865  GCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCTCCAGGAATGTAG   924

Qy   1026  CCATCCACCTTGTCAATCAAAGGGGGTTGGTGCTTCCGGACAT   1085
           |||||||||||||||||||||
Db    925  CCATCCCATCCTGTGCTAAATCAAAGGGGTTGGTCGCTTCCGGGATTGTTGTTGTTACATAACCGGACAT   984

Qy   1086  AATCATAGGACCTCTCACACACAGTTCGCCTCTTTTGATTAAACGGCCACCGGTTTTCCCGGT   1145
           |||||||||||||||||||||
Db    985  AATCATAGGACCTCTCACACACAGTTCGCCTCTTTTGATTAAACGGCCACCAGGTTTTCCCGGT   1044

Qy   1146  ATCCAGATCCACAA-CCTTCGCTTCAAAAAATGGAACAACTTTACCGGACCCCACCGGGTT   1204
           |||||||||||||||||||||
Db   1045  ATCCAGATCCACAAACCCTTCGCTTCAAAAAATGGAACAACTTTACCGGACCCCACCGGGTT   1104

Qy   1205  TATCATCCCCTCGGGTGTAATCAGAAATAGCTGAGTGTACTCTCAGTGACCCCATATCTT   1264
           |||||||||||||||||||||
Db   1105  TATCATCCCCTCGGGTGTAATCAGAAATAGCTGAGTGTACTCTCAGTGACCCCATATCTT   1164

Qy   1265  GCCTGATACCTGGCAGGATGGAACCTTCTTGGCCAACCGGCTTCCCGGACTTCCTTAGAGAGGG   1324
           |||||||||||||||||||||
Db   1165  GCCTGATACCTGGCAGATGGAACCTTCTTGGCCAACCGGCTTCCCGGACTTCCTTAGAGAGGG   1224

Qy   1325  GAGCGGCCACCAGAAGCAGCAATTTCGTGTAAATTAGATAAATCGTATTCGTTATTGCAGTGC   1384
           |||||||||||||||||||||
Db   1225  GAGCGGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTCGTTATTGCAGTGC   1284

Qy   1385  TTTTGGCGAAGAAGAAGAAATAGGGTTGGGTTGCACCAGCAGCGCGCAACTTTTGTATTCCT   1444
           |||||||||||||||||||||
Db   1285  TTTTGGCGAAGAAGAAGAAATAGGGTTGGGCCAGCAGCGCGCAACTTTTGTATTCCT   1344

Qy   1445  GAAAGGCTCCTCAGAGAAACAGCTCTTCTCTTCAAAATCTAATACAATTAAGACGACTTAGCCAC   1504
           |||||||||||||||||||||
Db   1345  GAAAGGCTCCTCAGAGAAACAGCTCTTCTCTTCAAATCTATACAATTAAGACGACTTAGCCAC   1404

Qy   1505  ATATCAAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGGAATGGAACAACTTA   1564
           |||||||||||||||||||||
Db   1405  ATATCAAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGGAACAACTTA   1464

Qy   1565  AAATGGCAGTATCCCGGAATCGGCAATGATTTGATTTGACTTGCAAAAATAGGAATC   1624
           |||||||||||||||||||||
Db   1465  AAATCCCAGTATCCCGGAATCGGCAATGATTTGATTTGACTTGCAAAAATAGGAATC   1524

Qy   1625  TCACCCAGGACAGTTCTATGAGGCAGCAGCGACACCCTTTAGGCAGCCAGTGAGATCCGAGAGG   1684
           |||||||||||||||||||||
Db   1525  TCACCCAGGACAGTTCTATGAGGCAGCAGCGACACCCTTTAGGCAGCCAGTGAGATCCAGAGG   1584

Qy   1685  AGTTCATGATCAGTGCAATTGTCCTTGTCCCTATCTGGAAGGACTCTGGCCACAAAATCGTATT   1744
           |||||||||||||||||||||
Db   1585  AGTTCATGATCAGTGCAATTGTCTTGTCCCTATCCGGAAGGACTCTGGCCACAAAATCGTATT   1644

Qy   1745  CATTAAAAACCGGGAGGTAGATGAGATGAGGTGTGACAGTGTACATCGACGA   1804
           |||||||||||||||||||||
Db   1645  CATTAAAAACCGGGAGGTAGATGAGATGAGGTGTGACGTGTACATCGACGA   1704

Qy   1805  AATCCGTTTTAGAATCCATGATAATAATTTTTTTTTTGGATGATTGGGGAGCTTTTTTTTGCACGT   1864
           |||||||||||||||||||||
Db   1705  AATCCGTTTTAGAATCCATGATAATAATTTTTTTGGATGATTGGGGAGCTTTTTTTTGCACGT   1764

Qy   1865  TCAAAAAATTTTTTGCAACCCCTTTTTGGAAACGAAACACACGGGTAGGCTGCCGAAAATGCCCA   1924
           |||||||||||||||||||||
Db   1765  TCAAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACACGGGTAGGCTGCCGAAAATGCCCA   1824

Qy   1925  TACTGTTGAGCAATTCACGTTCATTCATTATATAAATGTCGGTTCGGCGGGGGCCAACTGCAA   1984
           |||||||||||||||||||||
Db   1825  TACTGTTGAGCAATTCACGTTCATTCATTATATAAATGTCGGTTCGGCGGGGGCCAACTCCGA   1884

Qy   1985  TAAATAAACGGCCCCAACACCGGGCACATCAAAAGAATTGAAAGAGAGAGTTTTTCACTGCATAGGACGA   2044
           |||||||||||||||||||||
Db   1885  TAAATAAACGGCCCCAACACCGGGCATAAAAGAATTGAAAGAGAGAGTTTTTCACTGCATAGGACGA   1944

Qy   2045  TTCTCGTGATTTGTTATTCAGCGCCATATCGTTTCATAAGCTTCTGCCAACCGAACGGACATTT   2104
           |||||||||||||||||||||
Db   1945  TTCTGTGATTTGTTATTCAGCGCCATATCGTTTCATAAGCTTCTGCCAACCGAACGGACATTT   2004

Qy   2105  CGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCAATCTGTAAAAGCAATTGTTC   2164
           |||||||||||||||||||||
Db   2005  CGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCAATCTGTAAAAGCAATTGTTC   2064

Qy   2165  CAGGAACCAGGGCGTATCTTCTTTCATAGCCTTATGCAGTTGCTCTCCAGGCGGTTCCATCTT   2224
           |||||||||||||||||||||
Db   2065  CAGGAACCAGGGCGTATCTTCTTTCATAGCCTTATGCAGTTGCTCTCCAGGCGGTTCCATCTT   2124

Qy   2225  CCAGCGGGATAGAATGGGCGCGGGGGCCTTTCTTTATGTTTTGGGGCGTGCTTCTTCCATGG   2278
           |||||||||||||||||||||
Db   2125  CCAGCGGGATAGAATGGGCGCGGGGGCCTTTCTTTATGTTTTGGGGCGTCTTCTTCCATGG   2178
```

Search completed: September 29, 2005, 14:36:51
Job time : 11878 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 29, 2005, 11:18:45 ; Search time 1436 seconds
(without alignments)
11423.122 Million cell updates/sec

Title: US-10-734-801-18
Perfect score: 2771
Sequence: 1 ggatccgctgtggaatgtgt..........atgggaccacaaatctctc 2771

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4390206 seqs, 295987O667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2771	100.0	2771	6	ABQ78072	Abq78072 Partial p
2	2771	100.0	5860	6	ABQ78071	Abq78071 pMJ050 co
3	2674	96.5	5860	2	ABV45225	Aav45225 Human ner
4	2327	84.0	2327	6	ABQ78074	Abq78074 Partial p
5	1763.2	63.6	5064	5	AAA07776	Aaa07776 DNA seque
6	1763.2	63.6	5256	6	AAD27538	Aad27538 pGL3 enha
7	1763.2	63.6	7014	12	AAD27537	Aad27537 pGL3 cont
8	1763.2	63.5	7788	6	ADN11354	Adn11354 pcFUSII p
9	1758.2	63.6	5256	6	AAD27537	Aad27537 pGL3 cont
10	1747.2	63.1	7312	6	AAD31131	Aad31131 pcFUSII p
11	1663.4	60.0	11004	12	ADJ25519	Adj25519 ARE-LUC/C
12	1657.6	59.8	4818	6	AAD27536	Aad27536 pGL3 basi
13	1657.6	59.8	4987	3	AAA29136	Aaa29136 pNF-kappa
14	1657.6	59.8	4987	6	ABV73856	Abv73856 Plasmid p
15	1657.6	59.8	4987	6	ABS76596	Abs76596 Human ind
16	1657.6	59.8	5010	6	AAD27539	Aad27539 pGL3 prom
17	1657.6	59.8	5010	8	ACC69628	Acc69628 Cloning v
18	1657.6	59.8	5010	13	ADR72550	Adr72550 DNA of pG
19	1657.6	59.8	5105	11	ADO58341	Ado58341 Recombina
20	1657.6	59.8	5376	9	AAD55995	Aad55995 Luciferas
c 21	1657.6	59.8	5468	13	ADQ91240	Adq91240 Fusion co
c 22	1657.6	59.8	5924	13	ADS75093	Ads75093 Plasmid p
c 23	1657.6	59.8	5960	2	AAV45225	Aav45225 Human ner
c 24	1657.6	59.8	5982	13	ADS75092	Ads75092 Plasmid p
c 25	1657.6	59.8	8307	13	ADQ91241	Adq91241 Expressio
c 26	1657.6	59.8	8763	10	ACF04645	Acf04645 Vector co
c 27	1657.6	59.8	8771	10	ACF04644	Acf04644 Vector co
c 28	1657.6	59.8	15185	8	AAD52000	Aad52000 Human CYP
c 29	1657.6	59.8	21404	6	ABK86229	Abk86229 AIP-1/FLA
c 30	1656.6	59.8	7209	12	AAH24106	Aah24106 Mouse Per
c 31	1656.4	59.8	13654	4	AAD07501	Aad07501 pPRIBP+1u
c 32	1655.2	59.7	4413	12	ADN11342	Adn11342 Fluc-hygr
c 33	1655.2	59.7	7209	12	ADN11353	Adn11353 Fluc-GAGG
c 34	1655.2	59.7	14947	12	ADN11343	Adn11343 Targeting
c 35	1655	59.7	5143	4	AAD10006	Aad10006 Plasmid p
c 36	1655	59.7	5662	4	AAD10007	Aad10007 Plasmid p
c 37	1655	59.7	6310	4	AAD10005	Aad10005 Plasmid p
c 38	1655	59.7	12850	4	AAD07495	Aad07495 p2EO1P+1u
c 39	1654.4	59.7	4824	4	AAD10009	Aad10009 Plasmid p
c 40	1654.2	59.7	7460	4	AAH74867	Aah74867 Nucleotid
c 41	1654.2	59.7	14194	4	AAC66932	Aac66932 Plant sig
c 42	1653.4	59.7	1706	13	ADR12256	Adr12256 Luciferas
c 43	1653.4	59.7	2477	6	ABT07352	Abt07352 Reporter
c 44	1653.4	59.7	2477	10	ADF49175	Adf49175 Luciferas
c 45	1653.4	59.7	5475	8	ABV75872	Abv75872 Vector pT

ALIGNMENTS

RESULT 1
ABQ78072

ID   ABQ78072 standard; DNA; 2771 BP.
XX
AC   ABQ78072;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Partial pMJ050 construct SEQ ID NO 18.
XX
KW   Genomic replication; RNA-dependent RNA polymerase virus; RDRP virus;
KW   infection; ds.
XX
OS   Rhesus macaque polyoma virus.
OS   Hepatitis C virus.
OS   Hepatitis D virus.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key   Location/Qualifiers
FH
FT   promoter   1..347
FT   /*tag= a
FT   /note= "SV40_promoter"
FT   3'UTR   348..626
FT   /*tag= b
FT   /note= "Hepatitis C virus 3'UTR, antisense orientation"
FT   misc_feature   627..2284
FT   /*tag= c
FT   /note= "Luciferase sequence, antisense orientation"
FT   5'UTR   2285..2674
FT   /*tag= d
FT   /note= "Hepatitis C virus 5'UTR, antisense orientation"
FT   RBS   2675..2771
FT   /*tag= e
FT   /note= "Hepatitis delta virus ribozyme sequence, sense
FT   orientation"
XX
PN   WO200261048-A2.
XX
PD   08-AUG-2002.
XX
PF   31-JAN-2002; 2002WO-US0002952.
XX

PR 31-JAN-2001; 2001US-0265437P.
XX
PA (BRIS-) BRISTOL MYERS SQUIBB PHARMA CO.
XX
PI King RW, Jeffries MW, Pasquinelli C;
XX
DR WPI; 2002-619240/66.
XX
PT The invention relates to methods for measuring the genomic replication of RNA-dependent RNA polymerase (RDRP)
XX
PN Measuring the genomic replication of RNA-dependent RNA
CC virus, for designing therapies for the treatment of cells infected with
CC RDRP viruses, by transfecting cultured cells with a construct comprising
CC the cDNA.
CC
CC The invention relates to methods for measuring the genomic replication of
CC a virus that is dependent for replication upon RNA-dependent RNA
CC polymerase (RDRP) virus comprising transfecting cells with a construct
CC having, in antisense orientation, the cDNA of a reporter gene sequence
CC operably linked on its 5' end with the untranslated region (UTR) of the
CC native 3' end of the RDRP virus and operably linked on its 3' end with
CC the UTR of the native 5' end of the RDRP virus. The methods are useful
CC for designing therapies for the in vivo treatment of cells that are
CC infected with RDRP viruses. The methods are also useful to provide a
CC convenient platform for screening inhibitors to RDRP viral replication.
CC The present sequence is that of part of the pMJ050 construct (ABQ78071)
CC used in examples of the invention
XX
PS Claim 12; Fig 3; 60pp; English.
XX
SQ Sequence 2771 BP; 751 A; 700 C; 652 G; 668 T; 0 U; 0 Other;

Query Match   100.0%; Score 2771; DB 6; Length 2771;
Best Local Similarity   100.0%;
Matches 2771; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

[Sequence alignment blocks follow, with Qy (query) and Db (database) rows showing matching nucleotide sequences from position 1 to 2771]

```
Qy    1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCTATCGAGAGGACTCTGGCACAAAATCG 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCTATCGAGAGGACTCTGGCACAAAATCG 1740

Qy    1741 TATTCATTAAAACCGGGAGGTAGAGATGAGAGTGACGAAGCGTGTACATCGACTGAAATCCC 1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 TATTCATTAAAACCGGGAGGTAGAGATGAGAGTGACGAAGCGTGTACATCGACTGAAATCCC 1800

Qy    1801 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTTGC 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTTGC 1860

Qy    1861 ACGTTCAAAAATTTTTTCGAAACCCCTTTTTTGGAAAACGAACACCACGCGTAGGCTGCGAAATG 1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 ACGTTCAAAAATTTTTTCGAAACCCCTTTTTTGGAAAACGAACACCACGCGTAGGCTGCGAAATG 1920

Qy    1921 CCCATACTGTTGAGCAATTCACGTTCATTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACT 1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CCCATACTGTTGAGCAATTCACGTTCATTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACT 1980

Qy    1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040

Qy    2041 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATTAGACTTCTCTGCCAACCGGAAC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATTAGACTTCTCTGCCAACCGGAAC 2100

Qy    2101 ATTTCGAAGTACTCAGGCTAAGTAGTGATGTCCACCTCGATATATGTGTAAAAGCAATT 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 ATTTCGAAGTACTCAGGCTAAGTAGTGATGTCCACCTCGATATATGTGTAAAAGCAATT 2160

Qy    2161 GTTCCAGGAACCAGGCGTATCTCTTCATAGACGTTCATTCATCAGTTGCTCTCCAGCGGGTTCCA 2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 GTTCCAGGAACCAGGCGTATCTCTTCATAGACGTTCATTCATCAGTTGCTCTCCAGCGGGTTCCA 2220

Qy    2221 TCTTCCAGCGGATAGAATGGCGCCGGCCCTTCTTTATTGTTTTTTGGCGTCTTCCATGGGA 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 TCTTCCAGCGGATAGAATGGCGCCGGCCCTTCTTTATTGTTTTTTGGCGTCTTCCATGGGA 2280

Qy    2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTTGTGCTCATGATGCAC 2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTTGTGCTCATGATGCAC 2340

Qy    2341 GGTCTTACGAGAGACCTCCCGGGGGCACTCGGGCACCCTATCAGGCAGTACCACAAGGGCCTT 2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GGTCTTACGAGAGACCTCCCGGGGGCACTCGGGCACCCTATCAGGCAGTACCACAAGGGCCTT 2400

Qy    2401 TCGCGACCCACACTACTCTCGGCTAGCAGTCTTCGGGCGGGGCAGCGCCAAATCTTCCAGGCCAT 2460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 TCGCGACCCACACTACTCTCGGCTAGCAGTCTTCGGGCGGGGCAGCGCCAAATCTTCCAGGCCAT 2460

Qy    2461 TGAGCGGGGGTTATCCAAGAAAAGGACCCGGTCGTCCTCCGGCCAATTCCGGTGTACTCACCGGT 2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 TGAGCGGGGGTTATCCAAGAAAAGGACCCGGTCGTCCTCCGGCCAATTCCGGTGTACTCACCGGT 2520

Qy    2521 TCCGCAGACCACTATGGCTCTCCCGGGAGGCGGGGGTCTTGGAGGCTGCACGACACTCATA 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 TCCGCAGACCACTATGGCTCTCCCGGGAGGCGGGGGTCTTGGAGGCTGCACGACACTCATA 2580

Qy    2581 CTAACGCCATGGCTAGACGCTTTCTGCGTGAAGACAGTAGTTCCTCACAGGGGAGTGATT 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 CTAACGCCATGGCTAGACGCTTTCTGCGTGAAGACAGTAGTTCCTCACAGGGGAGTGATT 2640

Qy    2641 CATGGTGGAGTGTCGCGCCCCATCAGGGGCCATGGGTCCCAGCCTTCCTGCG 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 CATGGTGGAGTGTCGCGCCCCATCAGGGGCCATGGGTCCCAGCCTTCCTGCG 2700

Qy    2701 TGGCGCGCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTTAATGGCGAATGGGACC 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 TGGCGCGCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTTAATGGCGAATGGGACC 2760

Qy    2761 ACAAATCTCTC 2771
           |||||||||||
Db    2761 ACAAATCTCTC 2771
```

```
CC   for designing therapies for the in vivo treatment of cells that are
CC   infected with RDRP viruses. The methods are also useful to provide a
CC   convenient platform for screening inhibitors to RDRP viral replication.
CC   The present sequence is that of the pMJ0050 construct used in examples of
CC   the invention
XX
SQ   Sequence 5860 BP; 1465 A; 1518 C; 1464 G; 1413 T; 0 U; 0 Other;

Query Match          100.0%;  Score 2771;  DB 6;  Length 5860;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 2771;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

[DNA sequence alignment data: paired Qy (query) and Db (database) rows spanning positions 1–2040, with all positions matching (100% identity).]

Db 1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG

Qy 2041 ACGATTCGTGTGATTTGTATTCAGCCCATATCGTTTCTGCCAACCGAACGGAC 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 ACGATTCGTGTGATTTGTATTCAGCCCATATCGTTTCTGCCAACCGAACGGAC 2100

Qy 2101 ATTTCGAAGTACTCAGCGTAAGGTAGATGTTCCACCTCCGATATGTGCATCTGTAAAGCAATT 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 ATTTCGAAGTACTCAGCGTAAGGTAGATGTTCCACCTCCGATATGTGCATCTGTAAAGCAATT 2160

Qy 2161 GTTCCAGGAACCAGGGCGGTATCTCTTCATAGCCTTATGCAGTTGCTCTTCCAGCGGGTTCCA 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 GTTCCAGGAACCAGGGCGGTATCTCTTCATAGCCTTATGCAGTTGCTCTTCCAGCGGGTTCCA 2220

Qy 2221 TCTTCCACGGGGATAGAATGGGCGCCGGGGCCTTCTTCTTTATGTTTTTTTGGGGTCTTCCATGGGA 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 TCTTCCACGGGGATAGAATGGGCGCCGGGGCCTTCTTCTTTATGTTTTTTTGGGGTCTTCCATGGGA 2280

Qy 2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTGCTTCATGGATGCAC 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTGCTTCATGGATGCAC 2340

Qy 2341 GGTCTACGAGACCTCCGGGGCACTCGGAAGCACCTATCAGGCGAGTACCACAAGGCCTT 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 GGTCTACGAGACCTCCGGGGCACTCGGAAGCACCTATCAGGCGAGTACCACAAGGCCTT 2400

Qy 2401 TCCGACCCAAACACTACTCGGGCTAGCAGTCTTGCGGGGGCACGCGCCAAATCTCCAGGCAT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 TCCGACCCAAACACTACTCGGGCTAGCAGTCTTGCGGGGGCACGCGCCAAATCTCCAGGCAT 2460

Qy 2461 TGAGCGGGGGTTATCCAAGAAAGAACCCGGTCGTCCTGGCAATTCCGGTGTACTCACCGGT 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 TGAGCGGGGGTTATCCAAGAAAGAACCCGGTCGTCCTGGCAATTCCGGTGTACTCACCGGT 2520

Qy 2521 TCCGGACCACTATGGGCTTCTCCCGGGAGGCTGCACGACACTCATA 2580
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 TCCGGACCACTATGGGCTTCTCCCGGGAGGCTGCACGACACTCATA 2580

Qy 2581 CTAACGCCATGGCTAGAGACGGCTTTCTGCGTGAAGACAGTAGTTCCTCACAGGGGAGTGATT 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 CTAACGCCATGGCTAGAGACGGCTTTCTGCGTGAAGACAGTAGTTCCTCACAGGGGAGTGATT 2640

Qy 2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGCTGGCGGCCGGCATGGTCCCAGCCTTCCTCGC 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGCTGGCGGCCGGCATGGTCCCAGCCTTCCTCGC 2700

Qy 2701 TGGCGCCGGCTGGGCAACATTCCGAGGGGGACCGTCCCCTCGGTAATGGGGAATGGGACCC 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 TGGCGCCGGCTGGGCAACATTCCGAGGGGGACCGTCCCCTCGGTAATGGGGAATGGGACCC 2760

Qy 2761 ACAAATCTCTC 2771
         |||||||||||
Db 2761 ACAAATCTCTC 2771

RESULT 3

ABQ78073

ID   ABQ78073 standard; DNA; 2674 BP.
XX
AC   ABQ78073;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Partial pMJ050 construct SEQ ID NO 19.
XX
KW   Genomic replication; RNA-dependent RNA polymerase virus; RDRP virus;
KW   infection; ds.
XX
OS   Rhesus macaque polyoma virus.
OS   Hepatitis C virus.
OS   Unidentified.
OS   Chimeric.
XX

FH   Key   Location/Qualifiers
FH
FT   promoter   1..347
FT   /*tag= a
FT   /note= "SV40_promoter"
FT   3'UTR   348..626
FT   /*tag= b
FT   /note= "Hepatitis C virus 3'UTR, antisense orientation"
FT   misc_feature   627..2284
FT   /*tag= c
FT   /note= "Luciferase sequence, antisense orientation"
FT   5'UTR   2285..2674
FT   /*tag= d
FT   /note= "Hepatitas C virus 5'UTR, antisense orientation"
XX
PN   WO200261048-A2.
XX
PD   08-AUG-2002.
XX
PF   31-JAN-2002; 2002WO-US002952.
XX
PR   31-JAN-2001; 2001US-0265437P.
XX
PA   (BRIS-) BRISTOL MYERS SQUIBB PHARMA CO.
XX
PI   King RW, Jeffries MW, Pasquinelli C;
XX
DR   WPI; 2002-619240/66.
XX
PS   Example 1; Fig 3; 60pp; English.
XX
CC   The invention relates to methods for measuring the genomic replication of
CC   a virus that is dependent for replication upon RNA-dependent RNA
CC   polymerase (RDRP) virus comprising transfecting cells with a construct
CC   having, in antisense orientation, the cDNA of a reporter gene sequence
CC   operably linked on its 5' end with the untranslated region (UTR) of the
CC   native 3' end of the RDRP virus and operably linked on its 3' end with
CC   the UTR of the native 5' end of the RDRP virus. The methods are useful
CC   for designing therapies for the in vivo treatment of cells that are
CC   infected with RDRP viruses. The methods are also useful to provide a
CC   convenient platform for screening inhibitors to RDRP viral replication.
CC   The present sequence is that of part of the pMJ050 construct (ABQ78071)
CC   used in examples of the invention
XX
SQ   Sequence 2674 BP; 735 A; 665 C; 622 G; 652 T; 0 U; 0 Other;

Query Match   96.5%; Score 2674; DB 6; Length 2674;
Best Local Similarity 100.0%;
Matches 2674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCGGCTGTGGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCAGG 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GGATCCGGCTGTGGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCAGG 60

Qy 61 CAGAAGTATGCAAAAGCATGCAATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 CAGAAGTATGCAAAAGCATGCAATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120

Qy 121 CTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180

Qy 181 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA 240

Qy 241 TGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 TGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATT 300

```
Db  2461 TGAGCGGGGGTTATCCAAGAAAAGGACCCGGGTCGTCGTCCTGGCAATTCCGGTGTGTACTTCACCGGT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2521 TCCGCAGACCACTATGGCTTCTCCCGGGAGAGGCGGGGGGTCGTCGTCGGAGGCTGCACGCACACTCATA 2520

Db  2521 TCCGCAGACCACTATGGCTTCTCCCGGGAGAGGCGGGGGGTCGTCGTCGGAGGCTGCACGCACACTCATA 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2581 CTAACGCCATGGCTAGACCTTTCTGCGGTGAAGACAGTAGTTCCTCACAGGGGAGTGGATT 2580

Db  2581 CTAACGCCATGGCTAGACCTTTCTGCGGTGAAGACAGTAGTTCCTCACAGGGGAGTGGATT 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2581 CTAACGCCATGGCTAGACCTTTCTGCGGTGAAGACAGTAGTTCCTCACAGGGGAGTGGATT 2640

Qy  2641 CATGGTGGTGGAGTGTGTCGCCCCCATCAGGGGGGCTGGC 2674
         ||||||||||||||||||||||||||||||||||||||||
Db  2641 CATGGTGGTGGAGTGTGTCGCCCCCATCAGGGGGGCTGGC 2674
```

RESULT 4
ABQ78074

ID   ABQ78074 standard; DNA; 2327 BP.
XX
AC   ABQ78074;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Partial pMJ050 construct SEQ ID NO 20.
XX
KW   Genomic replication; RNA-dependent RNA polymerase virus; RDRP virus;
KW   infection; ds.
XX
OS   Hepatitis C virus.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key   Location/Qualifiers
FT
FT   3'UTR   1..279
FT   /*tag= b
FT   /note= "Hepatitis C virus 3'UTR, antisense orientation"
FT   misc_feature   280..1937
FT   /*tag= c
FT   /note= "Luciferase sequence, antisense orientation"
FT   5'UTR   1938..2327
FT   /*tag= d
FT   /note= "Hepatitis C virus 5'UTR, antisense orientation"
XX
PN   WO20026104B-A2.
XX
PD   08-AUG-2002.
XX
PF   31-JAN-2002; 2002WO-US002952.
XX
PR   31-JAN-2001; 2001US-0265437P.
XX
PA   (BRIS-) BRISTOL MYERS SQUIBB PHARMA CO.
XX
PI   King RW, Jeffries MW, Pasquinelli C;
XX
DR   WPI; 2002-619240/66.
XX
PT   Measuring the genomic replication of RNA-dependent RNA polymerase (RDRP)
PT   virus, for designing therapies for the treatment of cells infected with
PT   RDRP viruses, by transfecting cultured cells with a construct comprising
PT   the cDNA.
XX
PS   Example 1; Fig 3; 60pp; English.
XX
CC   The invention relates to methods for measuring the genomic replication of
CC   a virus that is dependent for replication upon RNA-dependent RNA
CC   polymerase (RDRP) virus comprising transfecting cells with a construct
CC   having, in antisense orientation, the cDNA of a reporter gene sequence
CC   operably linked on its 5' end with the untranslated region (UTR) of the
CC   native 3' end of the RDRP virus and operably linked on its 3' end with
CC   the UTR of the native 5' end of the RDRP virus. The methods are useful
CC   for designing therapies for the in vivo treatment of cells that are
CC   infected with RDRP viruses. The methods are also useful to provide a
CC   convenient platform for screening inhibitors to RDRP viral replication.
CC   The present sequence is that of part of the pMJ050 construct (ABQ78071)
CC   used in examples of the invention
XX
SQ   Sequence 2327 BP; 655 A; 562 C; 537 G; 573 T; 0 U; 0 Other;

Query Match   84.0%; Score 2327; DB 6; Length 2327;
Best Local Similarity 100.0%; Pred. No. 1.1e-302;
Matches 2327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  348 AGCTTACATGATCTTCCAGAGAGGGCCCAGTAATCAGCACTCTCTGCAGTCATCGCGGCTCACGG 407
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 AGCTTACATGATCTTGCAGAGAGGGCCCAGTAATCAGCACTCTCTGCAGTCATCGCGGCTCACGG 60

Qy  408 ACCTTTCACAGCTAGCCGTGACTAGGGCTTAAGATGGAGCCACCATTAAAGAGAAAGGAAGGAA 467
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ACCTTTCACAGCTAGCCGTGACTAGGGCTTAAGATGGAGCCACCATTAAAGAGAAAGGAAGGAA 120

Qy  468 AAGAAAGGAAAAAGGAAAAAGGAAAAGGAAAAAGGAAAAAGGAAAAAGGAAAAAAGAAAAA 527
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AAGAAAGGAAAAAGGAAAAAGGAAAAGGAAAAAGGAAAAAGGAAAAAGGAAAAAAGAAAAA 180

Qy  528 AAGAAAAGGAAAAAAGAAAAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGG 587
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AAGAAAAGGAAAAAAGAAAAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGG 240

Qy  588 AAAATGGGCTTAAGAGAGGGCCCGGAGTGGTTTAACCCCAACCTTTAAACGCGGAGATCTTTTCGCCCTT 647
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AAAATGGGCTTAAGAGAGGGCCCGGAGTGGTTTAACCCCAACCTTTAAACGCGGAGATCTTTTCGCCCTT 300

Qy  648 CTTGGCCTTTATGAGGGATCTCTCTGAATTTTTCTTGCGTCGGAGTTTTCCGGTAAGACCTTT 707
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CTTGGCCTTTATGAGGGATCTCTCTGAATTTTTCTTGCGTCGGAGTTTTCCGGTAAGACCTTT 360

Qy  708 CCGGTACTTCGTTCCACAACACAACTCCTCCGCGGCAAACTTTTTCGCGGTGTTGTTACTTGACT 767
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CCGGTACTTCGTTCCACAACACAACTCCTCCGCGGCAAACTTTTTCGCGGTGTTGTTACTTGACT 420

Qy  768 GGGCACGTAATCCACGATCTCTTTTCCGTCATCGTTCTTTCCGTGCTCCAAAAACAACAAC 827
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGGCACGTAATCCACGATCTCTTTTCCGTCATCGTTCTTTCCGTGCTCCAAAAACAACAAC 480

Qy  828 CTTAATCAGAGACTTCAGGCGGTCAAGGTCATCGTCGGGGAAGAAGTGTTCGTCTTCGTCCCAGTAAGC 887
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTTAATCAGAGACTTCAGGCGGTCAAGGTCATCGTCGGGGAAGAAGTGTTCGTCTTCGTCCCAGTAAGC 540

Qy  888 GTTGGGGTGTTGGAGCAAGATGGATTCCAATTCAGCGGAGCAAGCACCTGATAGCCTTTGTA 947
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GTTGGGGTGTTGGAGCAAGATGGATTCCAATTCAGCGGAGCAAGCACCTGATAGCCTTTGTA 600

Qy  948 CTTAATCAGAGACTTCAGGCGGTCGGGGGAGGAAGTTCACCGGGCGTCATCGTCGGGGAGGAT 1007
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAATCAGAGACTTCAGGCGGTCGGGGGAGGAAGTTCACCGGGCGTCATCGTCGGGGAGGAT 660

Qy 1008 TATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGGTTGGGCTCTTTGTA 1067
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGGTTGGGCTCTTTGTA 720

Qy 1068 GTTTACATAACCGGACAATAATCATTAGGAGCCTTCTCACACACAGTTCGCGCCCTTTGATTAAC 1127
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GTTTACATAACCGGACAATAATCATTAGGAGCCTTCTCACACACAGTTCGCGCCCTTTGATTAAC 780

Qy 1128 GCCCAGGCGTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTT 1187
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GCCCAGGCGTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTT 840

Qy 1188 ACCGACCGCGCCCCGGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGATTAGTCTC 1247
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 ACCGACCGCGCCCCGGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGATTAGTCTC 900

Qy 1248 AGTGAGCCCATATCCTTGCCTAGTACCCTGCAGATGGAACCTTCTTGGACAACCGCTTCCCC 1307
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   901 AGTGAGCCATATCCTTGCCTGATACCTGGCAACCTTCTTGGCAACGCTTCCCC  960
Qy       
Db  1308 GACTTCCTTAGAGAGGAGGGACCGGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTA 1367
Qy   961 GACTTCCTTAGAGAGGAGGGACCGGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTA 1020
Db  1368 TTTGTCAATCAGAGTGCTTTTGGCCAAGAAGGAGAGAATAGGGTTGGCACGCAGCAGCGCACT 1427
Qy  1021 TTTGTCAATCAGAGTGCTTTTGGCCAAGAAGGAGAGAATAGGGTTGGCACGCAGCAGCGCACT 1080
Db  1428 TTGAATCTTGTAATCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCTTCAAATCTTATACATTAA 1487
Qy  1081 TTGAATCTTGTAATCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCTTCAAATCTTATACATTAA 1140
Db  1488 GACGACTCGAAATCACATATCAAATATATCCGAGTCGTAGTCAGTAGTAGAAAACCGTGATG 1547
Qy  1141 GACGACTCGAAATCACATATCAAATATATCCGAGTCGTAGTCAGTAGTAGAAAACCGTGATG 1200
Db  1548 GAATGGAACAACACTTAAAATCGCAGTATCCGGAATGATTGATTGGCCAAAAAATAGGGATC 1607
Qy  1201 GAATGGAACAACACTTAAAATCGCAGTATCCGGAATGATTGATTGGCCAAAAAATAGGGATC 1260
Db  1608 TCTGGCATGCGGAGAATCTCACGCAGCAGCAGGCAGAGAGAGGACACCTTTAGGCAG 1667
Qy  1261 TCTGGCATGCGGAGAATCTCACGCAGCAGCAGGCAGAGAGAGGACACCTTTAGGCAG 1320
Db  1668 ACCAGTAGATCCAGAGGAGTTCATGAGTTCATTGCAAATTGTCTTGTCCCTATCGAAGGACTC 1727
Qy  1321 ACCAGTAGATCCAGAGGAGTTCATGAGTTCATTGCAAATTGTCTTGTCCCTATCGAAGGACTC 1380
Db  1728 TGGCACAAAATCGTATTCATTAAAAACCGGGGAGGTAGATGAGATGTGACCGAAGGTGTACAT 1787
Qy  1381 TGGCACAAAATCGTATTCATTAAAAACCGGGGAGGTAGATGAGATGTGACCGAAGGTGTACAT 1440
Db  1788 CGACTGAAATCCCTGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGG 1847
Qy  1441 CGACTGAAATCCCTGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGG 1500
Db  1848 GAGCTTTTTTGCACGTTCAAAAATTTTTTGCAACCCTTTTTGGAAACGAACACACACGGT 1907
Qy  1501 GAGCTTTTTTGCACGTTCAAAAATTTTTTGCAACCCTTTTTGGAAACGAACACACACGGT 1560
Db  1908 AGGCTGCGAAATGCCCATACTGTTGAGCAAATTCACGTTCATTATAAATGTCGTTCGCGGG 1967
Qy  1561 AGGCTGCGAAATGCCCATACTGTTGAGCAAATTCACGTTCATTATAAATGTCGTTCGCGGG 1620
Db  1968 CGCAACTGCAACTCCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTT 2027
Qy  1621 CGCAACTGCAACTCCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTT 1680
Db  2028 TTCACTGCATACGACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGC 2087
Qy  1681 TTCACTGCATACGACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGC 1740
Db  2088 CAACCGAACGGACGATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATC 2147
Qy  1741 CAACCGAACGGACGATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATC 1800
Db  2148 TGTAAAAGCAAATTGTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTC 2207
Qy  1801 TGTAAAAGCAAATTGTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTC 1860
Db  2208 TCCAGCGCGGTTCCATCTTCCAGCGGATAGAATGGCCGCCGGCCTTTCTTTATGTTTTTGGC 2267
Qy  1861 TCCAGCGCGGTTCCATCTTCCAGCGGATAGAATGGCCGCCGGCCTTTCTTTATGTTTTTGGC 1920
Db  2268 GTCTTCCATGGGGACGGTGGTTGGTTAGGTTTGGTTTTTCTTTGAGGTTTAGGATTCGT 2327
Qy  1921 GTCTTCCATGGGGACGGTGGTTGGTTAGGTTTGGTTTTTCTTTGAGGTTTAGGATTCGT 1980
Db  2328 GCTCATGATGCACGGTCTACGAGACCTCCCGGGGGCCACTCTGCAAGACACCCTATCAGGGCAGT 2387
Qy  1981 GCTCATGATGCACGGTCTACGAGACCTCCCGGGGGCCACTCTGCAAGACACCCTATCAGGGCAGT 2040
Qy  2388 ACCACAAAGGCCTTTCGCGACCCCAACACTACTCGGCTAGCAGTCTTCGCGGGGGGGGGGCCA 2447
Db  2041 ACCACAAAGGCCTTTCGCGACCCCAACACTACTCGGCTAGCAGTCTTCGCGGGGGGGGGGCCA 2100
Qy  2448 AATCTCCAGGCATTGAGCGCGGGGTTATCCAAGAAAGGACACCCGGTCAGCAATTCCGG 2507
Db  2101 AATCTCCAGGCATTGAGCGCGGGGTTATCCAAGAAAGGACACCCGGTCAGCAATTCCGG 2160
Qy  2508 TGTACTCACCGGTTCCGCAGACCACTATGGCTCCCGGAGGAGGCCTG 2567
Db  2161 TGTACTCACCGGTTCCGCAGACCACTATGGCTCCCGGAGGAGGCCTG 2220
Qy  2568 CACGACACTCATACTAACGCCATGGCTAGACGGCCTTTCTGCGTGAACAGTAGTTCCTCA 2627
Db  2221 CACGACACTCATACTAACGCCATGGCTAGACGGCCTTTCTGCGTGAACAGTAGTTCCTCA 2280
Qy  2628 CAGGGGAGTGATTCATGGTGGAGTGTCGCGCCCCATCAGGGGGGCTGGC 2674
Db  2281 CAGGGGAGTGATTCATGGTGGAGTGTCGCGCCCCATCAGGGGGGCTGGC 2327
```

CC   The patent discloses methods for determining the p53 status of a sample,
CC   useful for assaying for mimetics or antagonists of p53, or for the
CC   presence of DNA damage, comprises providing a sample containing a
CC   vector and determining whether p53 binds to the pGL3 vector. p53 is a
CC   transcription factor that regulates many genes including those associated
CC   with cell cycle control, apoptosis and DNA damage repair. pGL3 reporter
CC   vectors contain a modified firefly luciferase cDNA designated luc+. p53
CC   protein binds to pGL3-basic vector and causes luciferase expression. The
CC   method is useful for determining the p53 status of a sample. It is also
CC   useful for determining for mimetics or antagonists of p53 and for assaying
CC   for presence of activated p53 and/or DNA damage. The invention also
CC   relates to a method of modifying pGL3 vector which involves deletion or
CC   alteration of a CCCGG motif of the pGL3 vector and/or deleting or
CC   altering a sequence within 20 bp sequence 5' or 3' of CCCGG motif. The
CC   nucleic acid having a sequence incorporating the CCCGG motif is useful
CC   for conferring promoter activity or p53-responsiveness on a nucleic acid
CC   encoding a polypeptide of interest or in assays for p53 transcriptional
CC   activity. The present DNA sequence is pGL3 enhancer vector sequence
XX

SQ   Sequence 5064 BP; 1300 A; 1201 C; 1242 G; 1321 T; 0 U; 0 Other;

Query Match   63.6%; Score 1763.2; DB 6; Length 5064;
Best Local Similarity   87.9%; Pred. No. 2.5e-227;
Matches 2002; Conservative 0; Mismatches 168; Indels 108; Gaps 3;

```
Qy         1 GGATCCGCTGTGGAAATGTGTGTCAGTTAGGGTTGGGAAAGTCCCCAGCTCCCCAGCAGG 60
             ||||||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||
Db      2255 GGATCCGCTGTGGAAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 2196

Qy        61 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2195 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 2136

Qy       121 CTCCCCAGCGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCC 180
             |||||||| |||| |||||||||||||||||||||||||||||||||||||||||
Db      2135 CTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCC 2076

Qy       181 GCCCCTAACTCCCCGCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCTCCA 240
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      2075 GCCCCTAACTCCGCCCATCCCCCAGTTCCGCCCATTCTCCGCCTCCA 2016

Qy       241 TGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATT 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2015 TCG-------TTCAGATCCTTATCGATT 1995

Qy       301 CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTACATGATC 360
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db      1994 TTACCACATTTGTTAGAGAGTTTTACTTG----------------- 1968

Qy       361 TGCAGAGGAGGCCAGTATCAGCACTCTCCGAGTCATGCGGCGCTCACGGCGACCTTTCACAGCT 420
             |||| |||||||||||||||||||||||||||||||||||
Db      1967 ---------CTTTAAAAAACCTTCCACACTCCCCGAACCTGAAACATAAAAATGAA 1919

Qy       421 AGCCGGTGACTAGGGCTAAGATGGAGGCCACCATTAAAGAAGGAAGGAAGAAAGAAAA 480
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db      1918 TGCAAATTGGTTTGTTGTTTAACTTGTTTAATTGCAGCTTATAATATGGTTACAAATAACAATAG 1859

Qy       481 AGAAGGAAAGAAGAATAAAAAAAAAAAGGAAAAAGAAAAAAAAAAAAGAAAAGAAAAAAA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1858 CATCACAAAATTTCACAAATAAAGCCATTTTTTTCACTGCATTCTAGTTGTTGGTTTGTCAA 1799

Qy       541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAAATGGCCTAAGA 600
             |||||||||||||||||||||||||||||||||||||||||||
Db      1798 ACTCATCAATGTATCTTATCA--------------------------TGTCTGCTCGAAGC 1764

Qy       601 GGGCGGGAGTGGTTTACCCCGAACACTTTAAACGGCGGATCTCTTTTCCGCGCCCTTTCTTTATG 660
             |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1763 GGCCGGCGGCCGCCCCGGACTCTAGAATTACAGGCGGATCTCTTTCCGCCCCTTCTTTCTTTATG 1704

Qy       661 AGGATCTCTCTGATTTTTCTTGCGTCGACTTTTCCGGTCGAGCTCTTTCCGGTACTCGTCC 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1643 ACAAACAACTCCTTCACACAGTCGGGTTGTTCGGTTACTTGACTGGCGACGTTAATCC 1584

Qy       721 ACAAACACAACTCTCCTCCCGCAACTTTTTTCGCGTGTTGTTTAGTTGACTGGCGACCTAATCC 780
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1583 ACGATCTCTTTTCCGTCATCGTCTTTCCGGTCGAGCTCTTTCCGGTACTCGTCC 1524

Qy       781 ACGATCTCTTTTCCGTCATCGTCTTTCCGGTCGAGCTCTTTCCGGTACTCGTCC 840
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db      1523 TCACCGCGGGTCATCGTCGGGGAAGACCTGGCGACCAGCTGGTCGAAGATGTTTGG 1464

Qy       841 TCACCGCGGGTCATCGTCGGGGAAGACCTGGCGACCAGCTGGTCGAAGATGTTTGG 900
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      1463 AGCAAGATGGATTCCAATTCAGCGGGAGCCACCTTGACTAGCCTTTGTACTTAATCAGAGAC 1404

Qy       901 AGCAAGATGGATTCCAATTCAGCGGGAGCCACCTTGACTAGCCTTTGTACTTAATCAGAGAC 960
             ||||||||||||||||||||||||||||||||||||||||||
Db      1403 TTCAGGCGGTCAACGATGAAGAAGTGTTCGTCTTCGTGTCCCAGTAAGCTATGGTCTCCAGAA 1344

Qy       961 TTCAGGCGGTCAACGATGAAGAAGTGTTCGTCTTCGTGTCCCAGTAAGCTATGGTCTCCAGAA 1020
             |||||||||||||||||||||||||||||||||||||||||||||
Db      1343 TGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCG 1284

Qy      1021 TGTAGCCATCCATCCTTGTCAATCAAGGGCGTTGGTCGCTTCCGGATTGTTTACATAACCG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1283 GACATAATCATAGGACCTCTCACACAGTCGCCTCTTCGTTGGGGTTTTTC 1224

Qy      1081 GACATAATCATAGGACCTCTCACACAGTCGCCTCTTTGATTTAAACGGCGGACGTTTTTC 1140
             |||||||||||||||||||||||||||||||||||||||||||||
Db      1223 CCGGTATCCAGATCCACAACACTTGCTTCAGAATCCAACAACTTACCGGACGGCGCGCC 1164

Qy      1141 CCGGTATCCAGATCCACAACACTTGCTTCAGAAAAATGGAACAACTTACCGGACGCGCC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1163 GGTTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCGAGCCCATAT 1104

Qy      1201 GGTTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCGAGCCCATAT 1260
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      1103 CCTTGCCTGATACCTGGCAGATGAATGGAACCTCTTCCTTCCTTAGAG 1044

Qy      1261 CCTTGCCTGATACCTGGCAGATGAATGGGAACCTCTTCCTTCCTTAGAG 1320
             ||||||||||||||||||||||||||||||||||||||||||
Db      1043 AGGGGAGGCCGCCACCAGAAGCAAATTCGTGTAAATTAGATAAATCGTAAATTTGTCAATCAGA 984

Qy      1321 AGGGGAGGCCGCCACCAGAAGCAAATTCGTGTAAATTAGATAAATCGTAAATTTGTCAATCAGA 1380
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       983 GTGCTTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCGACTTGAAATCTTGTAA 924

Qy      1381 GTGCTTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCGCACTTGAAATCTTGTAA 1440
             |||||||||||||||||||||||||||||||||||||||||||||
Db       923 TCCTGAAGGGCTCCTCAGAAAACAGCTCTTTCTCAAATCTATACATTATACATTAAGAGACTCGAAAT 864

Qy      1441 TTACCACATTTGTTAGAGAGTTTTACTTG---------------- 1968
             |||
Db      1967 --------CTTTAAAAAACCTTCCACACTCCCCGAACCTGAAACATAAAAATGAA 1919
```

```
RESULT 6
ID   AAA07776 standard; DNA; 5256 BP.
XX
AC   AAA07776;
XX
DT   03-JUL-2000  (first entry)
XX
DE   DNA sequence of plasmid pGL2.
XX
KW   ced-6; h1ced-6; h2ced-6; signal transduction pathway; phagocytosis;
KW   cancer; autoimmune disease; neurodegenerative disease; stroke; AIDS;
KW   Huntington's disease; myocardial infarction; cytostatic; neuroprotective;
KW   cardiant; immunosuppressive; apoptosis modulator; luciferase; ss.
XX
OS   Synthetic.
XX
PN   WO9964586-A2.
PD   16-DEC-1999.
XX
PF   10-JUN-1999; 99WO-EP004043.
XX
PR   11-JUN-1998; 98GB-00012660.
PR   24-SEP-1998; 98GB-00020816.
XX
PA   (DEVG-) DEVGEN NV.
XX
PI   Smits E, Van Criekinge WMR, Bogaert TAOE;
XX
DR   WPI; 2000-246285/21.
XX
PT   Assays for determining the phagocytosis of apoptotic cells useful for
PT   identifying a compound which influences the phagocytic uptake of
PT   apoptotic cells and treats cancers and neurodegenerative diseases.
XX
PS   Example; Fig 19; 122pp; English.
XX
CC   The invention relates to assays involving two human homologues of
CC   Caenorhabditis elegans ced-6 (h1ced-6 and h2ced-6) for identifying
CC   compounds which function as an inhibitor or an enhancer of a signal
CC   transduction pathway. The assays are carried out by measuring
CC   phagocytosis of apoptotic cells. The methods are useful for identifying
CC   compounds which can act as apoptotic modulators which are useful for
CC   treating diseases such as cancer, autoimmune diseases, neurodegenerative
CC   diseases such as Huntington's disease, stroke, myocardial infarction and
CC   AIDS. The assays are well adapted for medium and high throughput
CC   screening using a multi-well plate format. The present sequence
CC   represents the DNA sequence of plasmid pGL2 which is suitable for
CC   introduction of reporter gene luciferase into Ba/K3 cells
XX
SQ   Sequence 5256 BP; 1336 A; 1268 C; 1281 G; 1371 T; 0 U; 0 Other;

Query Match             63.6%;  Score 1763.2;  DB 3;  Length 5256;
Best Local Similarity   87.9%;
Matches 2002;  Conservative    0;  Mismatches  168;  Indels  108;  Gaps  3;
```

```
Qy        1 GGATCCGCTGTGGAATGTGTTGCAGTTAGGTGTGGAAAGTCCCCCAGGCTCCCCAGCAGG       60
            ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db     2447 GGATCCGCTGTGGAATGTGTTCAGTTAGGTGTGGAAAGTCCCCCAGGCTCCCCAGCAGG   2388

Qy       61 CAGAAGTATCGCAAAGCCATGCAATCTCAATTAGTCAGCAACCACCAGGTGTGGAAAGTCCCCCAGG      120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2387 CAGAAGTATCGCAAAGCCATGCAATCTCAATTAGTCAGCAACCACCAGGTGTGGAAAGTCCCCCAGG   2328

Qy      121 CTCCCCGCGAGGCAGGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC      180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2327 CTCCCCGCGAGGCAGGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC   2268

Qy      181 GCCCCTAACTTCCGCCCATCCCGCCCCTAACTCCGCCCCAGTTCCGCCCCATTCTCCGCCCCA      240
            ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db     2267 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCCAGTTCCGCCCCATTCTCCGCCCCA   2208

Qy      241 TGGCTGACTAATTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATT      300
            |||                          |||
Db     2207 TCG------------------------------TTCAGATCCTTATCGATT   2187

Qy      301 CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTACATGATC      360
            |||||||||||||||||||||||||||||||||||||||||||||||
Db     2186 TTACCACATTTGTTAGAGGTTTTTACTTTG-------------------   2160

Qy      361 TGCAGAGAGCGCCAGTATCAGCACACTCTGCCAGTCATGCCGGCTCACGCGACCTTTCACAGCT      420
            ||||||||||||||||||||
Db     2159 --------------CTTTAAAAAACCTCCCACACTTCCCCCTGAACCTGAAAACATAAAATGAA   2111

Qy      421 AGCCGTGACTAGGGCTAAGATGAGCCACCATTAAAGAGGAAGGAGAAAGAAAGAAAAGAAAA      480
            ||||||||
Db     2110 TGCAAATTGTTGTTGTTTAACTTGTTATTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAG   2051

Qy      481 AGAGGAAAGAAGAAAGAAAAAAAAAAAAAAAGAAAGAAAGAAAGAAAGAAAAAGAAAA      540
            ||| ||
Db     2050 CATCACAAATTTCACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTTGGTTTGTGTCCAA   1991
```

```
Qy      543 TGGTAATCCGTTTAGAATCCATGATAATAATAATTTTTTGGATGATTGGATGGGAGCTTTTTTGC      563
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1990 TGGTAATCCGTTTAGAATCCATGATAATAATAATTTTTTGGATGATTGGATGGGAGCTTTTTTGC   1860 ... (cont.)

Qy     1501 CCACATATCAAATATCCGAGTGTAGTAGTAAAACATTCAAAAACCGTGATGGAATGGAACAACA    1560
            ||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db      863 CCACATATCAAATATCCGAGTGTAGTAGTAAACATTCAAAAACCGTGATGAATGGAACAACA      804

Qy     1561 CTTAAAATCGCAGTATCCGAGAATGATTTGATTGGCCAAAAAATAGGATCTCTGGCATGCGAG    1620
            |||||| ||||||||| ||| |||||||||||||||||||||||||||||||||||||||||
Db      803 CTTAAAAATCGCAGTACCCGGAATGATTTGATTGGCCAAAAAATAGGATCTCTGGCATGCGAG      744

Qy     1621 AATCTCACGCAGGCAGTTCTTATGAGGGCAGGACGAGACCAGTAGATCCA    1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      743 AATCTCACGCAGGCAGTTCTTATGAGGGCAGGACGAGACCAGTAGATCCA      684

Qy     1681 GAGGAGTTCATGATCAGTGCAATTGTCTCTTGTCTTCGGTCCATTCGGAAGGAGGACTCTGGCACAAAATCG    1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      683 GAGGAGTTCATGATCAGTGCAATTGTCTCTTGTCTTCGGTCCATTCGGAAGGAGGACTCTGGCACAAAATCG      624

Qy     1741 TATTCATTTAAAACCGGGAGGTAGAGATGAGAGTGAGACGGTTACATCGACTGAAATCCC    1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      623 TATTCATTTAAAACCGGGAGGTAGAGATGAGAGTGAGACGGTTACATCGACTGAAATCCC      564

Qy     1801 TGGTAATCCGTTTAGAATCCATGATAATAATTTTTTGGATGATTGGATGGGAGCTTTTTTGC    1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      563 TGGTAATCCGTTTAGAATCCATGATAATAATTTTTTGGATGATTGGATGGGAGCTTTTTTGC      504

Qy     1861 ACGTTCAAAATTTTTCGCACCCCTTTTTTTGGAAACGAACCACCACAGGGTAGGCTGGGAAATG    1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      503 ACGTTCAAAATTTTTGCAACCCCTTTTTTGGAAACGAACCACCACAGGGTAGGCTGGGAAATG      444

Qy     1921 CCCATACTGTTGAGCAATTCACGTTCATTATAATATAAAGTCGTTCGGCCACACCGTAGGCTGCAACTGCCAACT    1980
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      443 CCCATACTGTTGAGCAATTCACGTTCATTATAATATAAAGTCGTTCGGCCACACCGTAGGCTGCAACTGCCAACT      384

Qy     1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG    2040
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      383 CCGATAAATAACGCCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG      324

Qy     2041 ACGATTCTGTGATTTGTATTTCAGCGCCCCATAATCGTTTCAACCGAACGGAC    2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      323 ACGATTCTGTGATTTGTATTTCAGCGCCCCATATCGTTTCAACCGAACGGAC      264

Qy     2101 ATTTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGGATATGTGCATCTGTAAAAGCAATT    2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      263 ATTTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGGATATGTGCATCTGTAAAAGCAATT      204

Qy     2161 GTTCCAGGAACCAGGGCGTATCTCTTCATGAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA    2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      203 GTTCCAGGAACCAGGGCGTATCTCTTCATGAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA      144

Qy     2221 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGG    2278
            |||||||||||||||||||||||||||||||||||||||||||
Db      143 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGG       86
```

```
Qy    541 AAAAAAAAAAAAAAAAAAAAAAAACAGGGAAATGGCCTAAGA                     600
          |||                                    |||
Db   1990 ACTCATCAATGTATCTTATCA-----------------TGTCTGCTGCGAAGC         1956

Qy    601 GGCCGGGAGTGTTTACCCCAACCTTAAACGGGCGACCCTTCTTCTTGGCCTTTATG       660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1955 GGCCGGGCCGCCCGACTCAGGACGGCGACTCTTTCCGCCCTTCTTCTGGCCTTTATG     1896

Qy    661 AGGATCTCTCTGATTTTTCTTGCGGTGCGAGTTTCGGTAGCTACTTCGTCC           720
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1895 AGGATCTCTCTGATTTTTCTTGCGGTGCGAGTTTCGGTAGCTACTTCGTCC           1836

Qy    721 ACAAACACAAACTTCCTCCGCCCACTTTTTCTTGCCTCGGACGGTAATCCC           780
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1835 ACAAACACAAACTTCCTCCGCCCACTTTTTCTTGCCTCGGACGGTAATCCC           1776

Qy    781 ACGATCTCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACACAACACGG          840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1775 ACGATCTCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACACAACACGG          1716

Qy    841 TCACCGGCGGTCATGTGTCGGGAAGACCTGGGACACCTGGTCGGTCGGTAATCC        900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1715 TCACCGGCGTCATCGTCGGGGAAGACCTGGGACACCTGGTCGGTCGGTAATCC         1656

Qy    901 AGCAAGATGGATTCCAATTCAGGGGGGACCCACCTGATAGCCTTTGTTACTT          960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1655 AGCAAGATGGATTCCAATTCAGGGGGGACCCACCTGATAGCCTTTGTTACTT          1596

Qy    961 TTCAGGCGGTCAACGATGAAGAAGTTGTCTTCGTCTTCGTCGTAAGCTATGTCTCCAGAA   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1595 TTCAGGCGGTCAACGATGAAGAAGTTGTCTTCGTCTTCGTCGTAAGCTATGTCTCCAGAA   1536

Qy   1021 TGTAGCCATCCTCTTGTCAATCAAGGCGGTTGGTCGCTTCCGGATTGTTTACATAACG      1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1535 AGCAAGATGGATTCCATCCTTGTCAATCAAGGCGGTTGGTCGCTTCCGGATTGTTTACATAACG  1476

Qy   1081 GACATAATCATAGGACCTCTCCACACAGTCGCCTCTCTTGATTAATCAGA             1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1475 GACATAATCATAGGACCTCTCCACACAGTCGCCTCTCTTGATTAATCAGA             1416

Qy   1141 CCGGTATCCAGATCCACACAACCTTCGCTTCCTTCAAAAAAATGGAAACAACTTTAACCG   1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415 CCGGTATCCAGATCCACACAACCTTCGCTTCCTTCAAAAAAATGGAAACAACTTTAACCG   1356

Qy   1201 GGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATAT    1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1355 GGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATAT    1296

Qy   1261 CCTTGCCTGATACCTGGCAGATGGAACCTCTTCGGACAACCGCTTCCCTGACTTCCTTAGAG   1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1295 CCTTGCCTGATACCTGGCAGATGGAACCTCTTCGGACAACCGCTTCCCTGACTTCCTTAGAG   1236

Qy   1321 AGGGGAGGGCCACCAGAAGCAATTCGTGTAAATTAGATAAATCGTATTGTCAAATCAGA     1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1235 AGGGGAGGGCCACCAGAAGCAATTCGTGTAAATTAGATAAATCGTATTGTCAAATCAGA     1176

Qy   1381 GTGCTTTTGGCGAAGAAGGAGATAGGGTTGGCACCAGCAGCGCACTTGTTTGAATCTTGTAA   1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1175 GTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTGTTTGAATCTTGTAA   1116

Qy   1441 TCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTACATTAAGGACGACTCGAAAT   1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1115 TCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGGACGACTCGAAAT  1056

Qy   1501 CCACATATCAAAATATCCGAGTGTAGTAAACATTCCAAAACCGGTAGTGAATGAAACAACA   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1055 CCACATATCAAAATATCCGAGTGTAGTAAACATTCCAAAACCGGTAGTGAATGAAACAACA   996

Qy   1561 CTTAAAAATCGCAGTATCCGGAATGATTTGATTGCCAAAAAATAGGATCTCTCTGGCATTGCGAG   1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    995 CTTAAAAATCGCAGTATCCGGAATGATTTGATTGCCAAAAAATAGGATCTCTCTGGCATTGCGAG    936

Qy   1621 AATCTCACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTAGGCAGAGACCAGTAGATCCA   1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    935 AATCTCACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTAGGCAGAGACCAGTAGATCCA   876

Qy   1681 GAGGAGTTCATGATCAGTGCAAATTGTCTGTCTGTCCATTCATCGGAAGACTCG         1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    875 GAGGAGTTCATGATCAGTGCAAATTGTCTGTCTGTCCATTCATCGGAAGACTCG         816

Qy   1741 TATTCATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCC    1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    815 TATTCATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCC    756

Qy   1801 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTTGGTTTTTTTGGATGATAATGGAG    1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    755 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTTGGATGATAATGGTTTTTTTGC    696

Qy   1861 ACGTTCAAAAATTTTTTGCAACCCCTTTTTTGGAAACATAACGTGAACAACCACCGTCGCCAAAATG   1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    695 ACGTTCAAAAATTTTTTGCAACCCCTTTTTTGGAAACATAACGTGAACAACCACCGTCGCCGAAATG   636

Qy   1921 CCCATACTGTTGAGCAATTCACGTTCATTATTATAAATAAATGTCGTTCGCGGGTCGGGGTCGG   1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    635 CCCATACTGTTGAGCAATTCACGTTCATTATTATAAATAAATGTCGTTCGCGGGCGGGAACT   576

Qy   1981 CCGGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG   2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    575 CCGGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG   516

Qy   2041 ACGTTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGAC     2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    515 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGAC    456

Qy   2101 ATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATT   2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 ATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATT   396

Qy   2161 GTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA   2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    395 GTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA   336

Qy   2221 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGG    2278
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGG    278
```

RESULT 7

AAD27537/c

ID   AAD27537 standard; DNA; 5256 BP.
XX
AC   AAD27537;
XX
DT   18-APR-2002 (first entry)
XX
DE   pGL3 control vector DNA.
XX
KW   p53 protein; pGL3 luciferase reporter vector; luc+; transcription factor;
KW   cell cycle control; pGL3 control vector; apoptosis;
KW   firefly; ds.
XX
OS   Photinus pyralis.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key   Location/Qualifiers
FH
FT   misc_feature   1..41
FT   /*tag= a
FT   misc_feature   49..244
FT   /note= "Multiple cloning site"
FT   /*tag= b
FT   promoter   280..1929
FT   /*tag= b
FT   misc_feature   280..1929
FT   /note= "Luciferase gene (luc+)"
FT   /*tag= c
FT   primer_bind   complement(281..303)
FT   /*tag= d

```
FT   polyA_signal     /bound_moiety= "GL primer2"
FT                    1964..2185
FT                    /note= e
FT                    *tag=
FT   enhancer         2197..2441
FT                    /note= "SV40 late poly(A) signal"
FT                    *tag= f
FT   primer_bind      complement(2499..2518)
FT                    /tag= g
FT                    /bound_moiety= "RV primer4"
FT   misc_signal      2756
FT                    /tag= h
FT   misc_signal      complement(3521..4378)
FT                    /note= "ColE1-derived plasmid replication origin"
FT                    /tag= i
FT   misc_feature     4511..4965
FT                    /note= "Beta-lactamase gene"
FT                    /tag= j
FT   misc_signal      4965
FT                    /note= "F1 origin"
FT                    /tag= j
FT   polyA_signal     5096..5249
FT                    /tag= k
FT   primer_bind      5198..5217
FT                    /note= "Upstream poly(A) signal"
FT                    /tag= l
FT                    5217..1
FT                    /bound_moiety= "RV primer3"
XX
PN   WO200196602-A2.
XX
PD   20-DEC-2001.
XX
PF   18-JUN-2001; 2001WO-GB002718.
XX
PR   16-JUN-2000; 2000GB-0014820.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PI   Yang AL, Festing M;
XX
DR   WPI; 2002-130743/17.
XX
PS   Claim 20; Page 36-39; 53pp; English.
XX
CC   The patent discloses methods for determining the p53 status of a sample,
CC   useful for assaying for mimetics
CC   or antagonists of p53, or for the presence of DNA damage, comprises
CC   which comprise providing a sample containing a pGL3 luciferase reporter
CC   vector and determining whether p53 binds to the pGL3 vector. p53 is a
CC   transcription factor that regulates many genes including those associated
CC   with cell cycle control, apoptosis and DNA damage repair. pGL3 reporter
CC   vectors contain a modified firefly luciferase cDNA designated luc+. p53
CC   protein binds to pGL3-basic vector and causes luciferase expression. The
CC   method is useful for determining the p53 status of a sample. It is also
CC   useful for assaying for mimetics or antagonists of p53 and for assaying
CC   for presence of activated p53 and/or DNA damage. The invention also
CC   relates to a method of modifying pGL3 vector which involves deletion or
CC   alteration of a CCCGG motif of the pGL3 vector and/or deleting or
CC   altering a sequence within 20 bp sequence 5' or 3' of CCCGG motif. The
CC   nucleic acid having a sequence incorporating the CCCGG motif is useful
CC   for conferring promoter activity or p53-responsiveness on a nucleic acid
CC   encoding a polypeptide of interest or in assays for p53 transcriptional
CC   activity. The present DNA sequence is pGL3 control vector sequence
XX
SQ   Sequence 5256 BP; 1336 A; 1268 C; 1281 G; 1371 T; 0 U; 0 Other;

Query Match             63.6%;  Score 1763.2;  DB 6;  Length 5256;
Best Local Similarity   87.9%;
Matches  2002;  Conservative    0;  Mismatches  168;  Indels  108;  Gaps   3;
```

```
Qy          1 GGATCCGCGTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2447 GGATCCGCGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG 2388

Qy         61 CAGAAGTATGCAAAAGCAATGCATCTCAAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120
              |||||||||||||||||||| |||| |||||||||||||||||||||||||||||||||||||
Db       2387 CAGAACTATGCAAAAGCATGCAATTCTCAAATTACTCAGCAACCAGGTGTGGAAAGTCCCCAGG 2328

Qy        121 CTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATATGGTCCC 180
              |||||||||||||||||||| |||| ||||||||||||| ||||||||||||||||| ||||
Db       2327 CTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTTCCC 2268

Qy        181 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2267 GCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA 2208

Qy        241 TGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATT 300
              |                                                       |||
Db       2207 TCG-------------------------------------TTCAGATCCTTATCGATT 2187

Qy        301 CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTACATGATC 360
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2186 TTACCACATTTGTAGAGGTTTTACTTG----------------------------- 2160

Qy        361 TGCAGAGAGGCCAGTATCAGCACTCTCTGCAGTCATGCGGCGCTCACGGACCTTTCACAGCT 420
              ||                                                         ||
Db       2159 --------------CTTTAAAAAACCTCCCACACTCCCCCTGAAACATAAAAATGAA 2111

Qy        421 AGCCCTGACTAGGGCTAGGACCCACCATTAAAGAGGAGGAAGAAGAAGAAGAAAAA 480
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db       2110 TGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAAGGTTACATAAAGCAATAG 2051

Qy        481 AGAAGGAAAGAAGAAAAAAAGAAAAAAGAAAAAAAAAAAAGAAAAAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2050 CATCACAAATTTCACAAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTTCCA 1991

Qy        541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAAATGGCCTAAGA 600
              |||
Db       1990 ACTCATCAARTGTTACTTATCA--------------------------- 1956

Qy        601 GGCCGGAGTGTTACCCCAACCTTTTTTCGGGGTTTGTTACTTGACAGTTAATCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1955 GGCCGGCGGCCCCGACTCTAGAATTACACGGCGATCTTTTCGCGCCTTTCTTATG 1896

Qy        661 AGGGATCTCTTCTGATTTTTCCTTGCGTGCGGAGTTTTCCGGTAAGACCTTTCGTCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1895 AGGGATCTCTTCTGATTTTTCCTTGCGTGCGGAGTTTTCCGGTAAGACCTTTCGTCC 1836

Qy        721 ACAAACACAACTCCTCCGCGCCAACTTTTTCGCGGTTGTTACTTGACCGTAATCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1835 ACAAACACAACTCCTCCGCGCCAACTTTTTCGCGGTTGTTACTTGACCGTAATCC 1776

Qy        781 ACGATCTCTTTTTTCCGTCATCGTCTTTTCCGGTGCCTCCAAACACAACGCGGGGAAGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1775 ACGATCTCTTTTTCGTCATCGTCTTTTCCGGTGCCTCCAAACACAACGCGGGGAAGT 1716

Qy        841 TCACCCGGCGGCGCTTCATCGTCGGGGGAAGACCTGGAGCACACTTGCGGTGGGAAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1715 TCACCCGGCGCTGCATCGTCGGGGAAGACCTGGAAGCACACTGCGGTGGGGAAGT 1656

Qy        901 AGCAGAGATGGATTCCAAATTCAGCGGGAGCCACCTGAATAGCCGTTTGTTACTTCAGAGAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1655 AGCAAGGATGGATTCCAAATTCAGCGGGAGCCACCTGTTGTTACTTAATCAGAGAC 1596

Qy        961 TTCAGGGCGTCAACGGATGAGAGAACTGTTTGCTGTCTTCTGTCTTTGTTCGGGAA 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1595 TTCAGGGCGGTCAACGATGAGAAGAGTGTTCGTCTTCTGTCTTTGTTACGGAA 1536

Qy       1021 TGTAGCCATCCATCCTTGTCAATCAGGGGTTGGTTCGCTTCCGGAATTGTTACATAACCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1535 TGTAGCCATCCATCCTTGTCAATCAAGGGGTTGGTTCGCTTCCGGATTGTTACATAACCG 1476

Qy       1081 GACATAATCATTAGGAACCTCACACAGTTGCGCCTTCTCACACACAGTTCGCCTCTTCTTTTTC 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1475 GACATAATCATTAGGAACCTCTCACACAGTTGCGCCTTCTCACACACAGTTCGCCTCTTCTTTTTC 1416
```

```
Qy   1141 CCCGGTATCCAGATCCACAACCTTCGGCTTCAAAAAAATGGAACAACTTTACCGACCGCGCCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1415 CCCGGTATCCAGATCCACAACCTTCGGCTTCAAAAAAATGGAACAACTTTACCGACCGCGCCC 1356

Qy   1201 GGTTTATCATCCCCTCGGGTGTTAATCAGAGAATAGCTGATGTTAGTCTCAGTGAGCCCATAT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1355 GGTTTATCATCCCCTCGGGTGTTAATCAGAGAATAGCTGATGTTAGTCTCAGTGAGCCCATAT 1296

Qy   1261 CCTTGCCTGATACCTGGCAGGATGGAAACCTCTTGGCAACCGCTTCCCCGGACTTCCTTAGG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1295 CCTTGCCTGATACCTGGCAGGATGGAAACCTCTTGGCAACCGCTTCCCCGGACTTCCTTAGAG 1236

Qy   1321 AGGGGAGCGGCCCACCAGAAGCCAATTCGTGTAAATTAGAAATACGTATTTTGTCAAATCAGA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1235 AGGGGAGCGGCCCACCAGAAGCCAATTCGTGTAAATTAGATAAATCGTATTTTGTCAAATCAGA 1176

Qy   1381 GTGGCTTTTGGCGGAAGCAGAATAGGGGTTGGGCACCAGCAGCGGGCCACTTTGAATCTTGTAA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1175 GTGGCTTTTGGCGGAAGCAGAATAGGGGTTGGGCACCAGCAGCGGGCCACTTTGAATCTTGTAA 1116

Qy   1441 TCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCTTCAAAATCTATACATTAAGACGACTCGAAAT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1115 TCCTGAAGGCTCCTCAGAAAACAGCTCTTCTTCTTCAAAATCTATACATTAAGACGACTCGAAAT 1056

Qy   1501 CCACATATCAAAATATCCCAGTGTAGTGAAACATTCCAAAAACCGTGATGAAATGGAACAACA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1055 CCACATATCAAAATATCCCAGTGTAGTGAAACATTCCAAAAACCGTGATGAAATGGAACAACA 996

Qy   1561 CTTAAAATGCAGTATCCGGAATCGATTTGATTGCCAAAATAGGATCTCTGGAATCTTGTAA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   995  CTTAAAATGCGACTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGAATCTTGTAA 936

Qy   1621 AATCTCACCCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCCAGAGACCAGTAGATCCA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   935  AATCTCACCCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCCAGAGACCAGTAGATCCA 876

Qy   1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCGGCACAAAAATCG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   875  GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCGGCACAAAAATCG 816

Qy   1741 TATTCATTAAAACCGGGAGTAGATGAGATGAGATGAGAATGAGAATGGACGAACGTGTACATCGACTGAAATCCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   815  TATTCATTAAAACCGGGAGTAGATGAGATGAGATGAGAATGAGAATGGACGAACGTGTACATCGACTGAAATCCC 756

Qy   1801 TGGTAATCCGTTTGAATCCATGATAATATAATTTTTTTGGATGATTGGGAGCTTTTTTTTTGC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   755  TGGTAATCCGTTTGAATCCATGATAATATAATTTTTTTGGATGATTGGGAGCTTTTTTTTTGC 696

Qy   1861 ACGTTCAAAAATTTTTTGCAAACCCCTTTTTTGGAAAGACGAACACACACGGTAGGCTCGGAAATG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695  ACGTTCAAAAATTTTTTGCAAACCCCTTTTTTGGAAAGACGAACACACACGGTAGGCTCGGAAATG 636

Qy   1921 CCCATACTGTTGAGCAATTCAGTTCAGTTCATTATTATAAATTGCGTTCGGTCCGGGGCCAACTGCCAACT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   635  CCCATACTGTTGAGCAATTCAGTTCATTATTATAAATGTCGTTCAGTGCGTTCGGTCCGGGGCCAACTGCCAACT 576

Qy   1981 CCGATAAAATAAACGGGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   575  CCGATAAAATAAACGGGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 516

Qy   2041 ACGATTCTGTGATTTGTATTCAGCGCTTCTTCATAGCTTCTGCCAACCGAACGGAC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   515  ACGATTCTGTGATTTGTATTCAGCGCTTCTTCATAGCTTCTGCCAACCGAACGGAC 456

Qy   2101 ATTTCGAAGTACTCAGCGTAAGTAGTGATGTCCACCTCGATATGTCGAATCTGTAAAAGCAATT
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   455  ATTTCGAAGTACTCAGCGTAAGTAGTGATGTCCACCTCGATATGTCGAATCTGTAAAAGCAATT 396

Qy   2161 GTTCCAGGAACCAGGCGGTATCTCTTCAGTGATGTCATCTGTTCTCCAGCCGTTCCA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   395  GTTCCAGGAACCAGGCGGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 336

Qy   2221 TCTTCCCAGCGCGGATAGAATGGCGCGCGGGGCCTTTCTTTATGTTTTTGGGGTCTTCCATGG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2278 TCTTCCAGCGGGATAGAATGGCGCGGGGCCTTTCTTTATGTTTTTGGGGTCTTCCATGG 278
```

Db   335 TCTTCCAGCGGGATAGAATGGCGCGGGGCCTTTCTTTATGTTTTTGGGGTCTTCCATGG 278

RESULT 8

```
ADN11354/c

ID   ADN11354 standard; DNA; 7014 BP.
XX
AC   ADN11354;
XX
DT   15-JUL-2004  (first entry)
XX
DE   GAGGS-Fluc nucleotide sequence SEQ ID NO:15.
XX
KW   expression vector; short hairpin RNA; shRNA;
KW   expression vector; gene knock down; inducible gene knock down;
KW   constitutive gene knock down; inducible gene knock down;
KW   polymerase II dependent locus; gene therapy; gene; ds.
XX
OS   Synthetic.
XX
PN   WO200403578 2-A2.
XX
PD   29-APR-2004.
XX
PF   10-OCT-2003; 2003WO-EP011233.
XX
PR   17-OCT-2002; 2002EP-00023283.
PR   02-MAY-2003; 2003US-0467814P.
PR   10-JUL-2003; 2003US-0485969P.
XX
PA   (ARTE-) ARTEMIS PHARM GMBH.
XX
PI   Seibler J, Schwenk F, Kuehn R, Kueter-Luks B;
XX
DR   WPI; 2004-348459/32.
XX
PT   New expression vector comprising a short hairpin RNA (shRNA) construct
PT   under the control of a ubiquitous promoter, useful for constitutive
PT   and/or inducible gene knock down in a vertebrate, or in a tissue.
XX
PS   Example 3; SEQ ID NO 15; 70pp; English.
XX
CC   The present invention describes an expression vector comprising a short
CC   hairpin RNA (shRNA) construct under the control of a ubiquitous promoter.
CC   Also described: (1) a method for constitutive and/or inducible gene knock
CC   down in a vertebrate, or in a tissue culture or cells of a cell culture
CC   derived from a vertebrate, by stably integrating an expression vector into
CC   the genome of the vertebrate, of the tissue culture or of the cells of
CC   the cell culture; and (2) a vertebrate or tissue or cell culture derived
CC   from a vertebrate having stably integrated, preferably at a polymerase II
CC   dependent locus or the vertebrate, tissue culture or cells of the tissue
CC   culture, an expression vector as defined above. The expression vector can
CC   be used in gene therapy. The expression vector is useful for preparing an
CC   agent, shRNA, for constitutive and/or inducible gene knock down in a
CC   vertebrate, or in a tissue culture or cells of a cell culture derived
CC   from a vertebrate. The present sequence represents a nucleotide sequence
CC   which is used in the exemplification of the present invention.
XX
SQ   Sequence 7014 BP; 1588 A; 1756 C; 1953 G; 1717 T; 0 U; 0 Other;
```

Query Match   63.6%; Score 1763.2; DB 12; Length 7014;
Best Local Similarity   87.9%; Pred. No.2.4e-227;
Matches 2002; Conservative 0; Mismatches 168; Indels 108; Gaps 3;

```
Qy      1 GGATCGGCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCCAGGCAGCAGG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4183 GGATCGGCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCCAGGCAGCAGG 4124

Qy     61 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4123 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCC 4064

Qy    121 CTCCCCAGCGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 9
AAD31131

```
Db    4063 CTCCCCAGCAGCAGAAGTATGCAAAAGCATCTCAATTAGTCAGCAACCATAGTCCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    3091 GGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGATGGTAGTGAGCCCCATAT 3032

Db     181 GCCCCTTAACTCTCCGCGCCCACTCCCGCCCCTAACTCGTTCCCGCGCCCAGTTCCGCCCCCAGTTCAGCACCATAGTCAC
Qy    1261 CCTTGCCTGATACCTTCCTGCAGATGGAAACCTCTTTGGCAAACCTTCCCGCACTTCCTTAGAG 1320

Db    4003 GCCCCTTAACTCTCCGCGCCCATCCCGCCCCATTCTCCCGCTCCA 3944
Qy    3031 CCCTTGCCTGATACCTGGTCAGATGGAAATACCTCTTGGCAAACCCTCCTTAGAG 2972

Db     241 TGGCTGACTAATTTTTTTTTTATTATGCAGAGGCCCGGCCCTCGGCCTTGAGCTATT
Qy    1321 AGGGGAGGCGCCACCAGAGAAGCAATTTGCTGTAAATTGAATACGTTAATCAGA 1380

Db    3943 TCG
Qy    2971 AGGGGAGGCGCCACCAGAGAAGCAATTCGTGTAAATTAGATAAATCGTATTGTCAATCAGA 2912

Db     301 CCAGAAGTAGTGAGGAGGCCTTTTTGGGAGGCCTAGGCTTTTCAAAAAGTTACATGATC
Qy    1381 GTGCTTTTGGGCGAAGGAAGAGAAATAGGGTTGGCACGCACGCACTTTGAATCTTGTAA 1440

Db    3922 TTACCACATTTGTAGAGGTTTTTACTTTG
Qy    2911 GTGCTTTTGGGGGAAGGAAGAGAAATAGGGTTGGCACGCACGCAGTTGAATCTTGTAA 2852

Db     361 TGCAGAGAGGCCAGTATCAGCACTCTCTGCAGTCATGCGGCTCACGGACCTTTCACAGCT
Qy    1441 TCCTGAAAGGGCTCCTCAGAAAACAGCTCTTCTTCTTCAAAATCTATACATTAAAGACTCCA 1500

Db    3895 CTTTAAAAAACCTCCCACACACCCTGAAACCATAAAAATGAA
Qy    2851 TCCTGAAGGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAAGACTCCAAAT 2792

Db     421 AGCCGTGACTTAGGGCCTAAGATGGGAGCCACCATTAAAGAAGGAAGAGAAGAGAAAAA
Qy    1501 CCACATATCAAAATATCCGACTGGTAGTAACATTCCAAAACCGTGATGGGAATGGGAACAACA 1560

Db    3846 TGCAATTGTTGTTGTTAATCTTGTTTATTGCAGCTTATACAAATAAAAAGCAATAG
Qy    2791 CCACATATCAAAATATCCGACTGTAGTAACATTCCAAAACCGTGATGGGAATGGGAACAACA 2732

Db     481 AGAAGGAAGAGAAAGGAAAAAAAAGAAAAAAAAAAAAAAGAAAAAAAA
Qy    1561 CTTAAAATCGGAAGTAATCCGGAATGAATTGTTGATTGGCCAAAAATAGATGGCATGCGAG 1620

Db    3786 CATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCAATTCTAGTGTGTGGTTTGTTCCAA 3727
Qy    2731 CTTAAAAATCGGACGAGTAATCCGGAATGATTTTGATTGGCCAAAAATAGATGGCATGCGAG 2672

Db     541 AAAAAAAAAAAAAAAAAAAAAAAAACAGGAAATGGCCTAAGA
Qy    1621 AATCTCACGCAGGCAGTCTTATGACGGCAGAGCGACACCTTTAGGCAGAACCAGTAGATCCA 1680

Db    3726 ACTCATCAAAGTATCTTATCA
Qy    2671 AATCTCACGCAGGCAGTCTTATGACGGCAGAGCGACACCTTTAGGCAGAACCAGTAGATCCA 2612

Db     601 GGGCCGGAGTGTTACCCAACCTTAAACGGCGATCTTTCCGCCCCTTCTTGGCCTTTATG
Qy    1681 GAGGAGTTCATGACGTGCAATTGTTCTTGTCCCTATCGGAAGGACTCTGGGCACAAATCG 1740

Db    3691 GGCCGGCGCCGCCCGACTCTAGAAATTACACGGCGGATCTTTCGCCCTTCTTGGCCTTTATG
Qy    2611 GAGGAGTTCATGACGTGCAATTGTTCTTGTCCCTATCGGAAGGACTCTGGGCACAAATCG 2552

Db     661 AGGATCTCTCTGATTTTTCTTGCGTCGAGATTTTTCGGGGACGGTAATCC
Qy    1741 TATTCATTAAAACCGGGAGGTAGATGAGTGACGGAACGTGTACATCGACTGAAAATCCC 1800

Db    3631 AGGATCTCTCTGATTTTCTTGCGTCGAGATTTTTCGGGGACGGTAATCC
Qy    2551 TATTCATTAAAACCGGGAGGTAGATGAGTGACGGAACGTGTACATCGACTGAAAATCCC 2492

Db     721 ACAAACACAACTCCTCCCGGCCAACTTTTTTCGGCGTTGTTACTTGACTGGGGACCGTAATCC
Qy    1801 TGGTAAATCCGTTTTAGAAATCCATAGATAATAAAATTTTTTGGAAGCAGAACGTGACAGCTT 1860

Db    3571 ACAAACACAACTCCTCCCGGCCAACTTTTTCGGCGTTGTTACTTGACTGGGGACCGTAATCC 3512
Qy    2491 TGGTAAATCCGTTTTAGAAATCCATAGATAATAAAATTTTTTGGAAGCAGAACGTGACAGCTT 2432

Db     781 ACGAACTCTCTTTTCCGCTCATCGTCTTTCCGTTGTTCGTCCGGTAAGCTAGTGTCTCCAGAA
Qy    1861 ACGTTCAAAAATTTGATTTGTTATTCAGCCCATAGCCTTCATAGCTTCTGCCAAACCGAACGGAC 1920

Db    3511 ACGAACTCTCTTTTCCGCTCATCGTCTTTCCGTTGTTCGTCCGGTAAGCTAGTGTCTCCAGAA 3452
Qy    2431 ACGTTCAAAAATTTGATTTGTTATTCAGCCCATAGCCTTCATAGCTTCTGCCAAACCGAACGGAC 2372

Db     841 TCACCGGGCGTCATCGTGTCGGGAAGACCTGCCACACCTGCTCCAAAAACAACAACGGGCGGGGAAGT
Qy    1921 CCCATACTGTTGAGCAATTCACGTTCATTATTATATTTCATTCGTTCGCGGCGGCGGCGGCAACTCCAACT 1980

Db    3451 TCACCGGGCGTCATCGTGTCGGGAAGACCTGCCACACCTGCTCCAAAAACAACAACGGGCGGGGAAGT 3392
Qy    2371 CCCATACTGTTGAGCAATTCACGTTCATTATTATATTTCATTCGTTCGCGGCGGCGGCGGCAACTCCAACT 2312

Db     901 AGCAAGATGGATTCCAATTCAGCGGAGCCACCTGATAGGCTTGTTGTACTTAATCAGAGAC
Qy    1981 CCGATAAAATAAACGGCCCCAACACACGGGCATAAAGAATTGAAGACGAGTTTTCACTCGCATACG 1920

Db    3391 AGCAAGATGGATTCCAATTCAGCGGAGCCACCTGATAGGCTTGTTGTACTTAATCAGAGAC 3332
Qy    2311 CCGATAAAATAAACGGCCCCAACACACGGGCATAAAGAATTGAAGACGAGTTTTCACTCGCATACG 2252

Db     961 TTCAGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTTCGTCCGGTAAGCTAGTCTCCAGAA
Qy    2041 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCATAGCTTCTGCCAACCGGAC 2100

Db    3331 TTCAGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTTCGTCCGGTAAGCTAGTCTCCAGAA 3272
Qy    2251 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCATAGCTTCTGCCAACCGGAC 2192

Db    1021 TGTAGCCATCCATCCTTCTGTCAATCAAAGGGGTTGGTCGCCTTCCGATTTGTTACATAACCG
Qy    2101 ATTTCGAAGTACTCAGCGTAAGTGAGTGCAGTCGATAATGTGCATCTGTAAAAGCAATT 2160

Db    3271 TGTAGCCATCCATCCTTGTCAATCAAAGGGGTTGGTCGCCTTCCGATTTGTTACATAACCG 3212
Qy    2191 ATTTCGAAGTACTCAGCGTAAGTGAGTGCAGTCGATAATGTGCATCTGTAAAAGCAATT 2132

Db    1081 GACATAAATCATAGAGACCTCTCACACACAGTTCGGCCTTCTCTTTGATTAAACGGCCCAGGGTTTTTC
Qy    2161 GTTCCAGGAACCACCAGGCGGTAATCTCTTCATAGCGGTAAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2220

Db    3211 GACATAAATCATAGAGACCTCTCACACACAGTTCGGCCTTCTCTTTGATTAAACGGCCCAGGGTTTTTC 3152
Qy    2131 GTTCCAGGAACCACCAGGCGGTAATCTCTTCATAGCGGTAAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2072

Db    1141 CCGGTATCCAGATCCAACAACTTGCTTCAAAAAAATGGAACAACTTTACCGACGAGCCGCCC
Qy    2221 TCTTCCAGCGGGATAGAAATGGCGCCGGGCCCCGGGCCTTTCTTTAGTGTTTTTGGGCGTCTTCCATGG 2278

Db    3151 CCGGTATCCAGATCCAACAACTTGCTTCAAAAAAATGGAACAACTTTACCGACGAGCCGCCC 3092
Qy    2071 TCTTCCAGCGGGATAGAAATGGCGCCGGGCCCCGGGCCTTTCTTTAGTGTTTTTGGGCGTCTTCCATGG 2014

Qy    1201 GGTTTATCATCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTGTCTCAGTGAGCCCCATAT 1260
```

```
ID   AAD31131 standard; DNA; 7788 BP.
XX
AC   AAD31131;
XX
DT   21-AUG-2002  (first entry)
XX
DE   pcFUSII plasmid.
XX
KW   Reporter construct; transcription control element; cell surface receptor;
KW   G-protein coupled receptor; GPCR; tyrosine kinase-type receptor;
KW   ion channel; high-throughput screening; HTS; jelly fish; EGFP;
KW   enhanced green fluorescent protein; TRE; TPA responsive element;
KW   12-0-tetradecanoylphorbol-13-acetate; CMV; promoter; firefly; luciferase;
KW   simian virus 40; SV40; pcFUSII plasmid; chimeric; ds.
XX
OS   Cytomegalovirus.
OS   Aequorea victoria.
OS   Photinus pyralis.
OS   Rheus macaque polyoma virus.
OS   Synthetic.
OS   Chimeric.
XX
PN   WO20220749-A2.
XX
PD   14-MAR-2002.
XX
PF   06-SEP-2001; 2001WO-IB001938.
XX
PR   07-SEP-2000; 2000US-0230705P.
XX
PA   (OWMA/) OWMAN C S O.
PA   (OLDE/) OLDE B A.
PA   (KOTA/) KOTARSKY K.
XX
PI   Owman CSO,  Olde BA,  Kotarsky K;
XX
DR   WPI; 2002-415727/44.
XX
PT   Reporter construct with a chimeric reporter gene linked to transcription
PT   control element(s), useful for detecting substances that interact with
PT   cell surface receptors, e.g. G-protein coupled receptor family or ion
PT   channels.
XX
PS   Example 2; Page 59-62; 66pp; English.
XX
CC   The invention relates to a reporter construct, comprising a chimeric
CC   reporter gene (comprising coding sequences from two different genes fused
CC   to produce a gene product that is detectable without the need to lyse or
CC   otherwise destroy or diminish the viability of the cell in which they are
CC   expressed) operably linked to at least one transcription control element.
CC   The reporter construct is useful for detecting substances that interact
CC   with cell surface receptors, such as those of the G-protein coupled
CC   receptor family, tyrosine kinase-type receptors or ion channels. The
CC   construct is particularly useful in high-throughput screening assays. The
CC   construct is used in recombinant G-protein coupled receptor assays which
CC   are more sensitive and less labour and time-intensive than previous
CC   assays. The present sequence is pcFUSII plasmid comprising TPA (12-0-
CC   tetradecanoylphorbol-13-acetate) responsive elements (TREs), minimal
CC   Cytomegalovirus promoter, Aequorea victoria enhanced green fluorescent
CC   protein (EGFP) gene, firefly luciferase gene, simian virus 40 (SV40)
CC   promoter and neomycin resistance cassette
XX
SQ   Sequence 7788 BP; 1777 A; 2051 C; 1970 G; 1990 T; 0 U; 0 Other;

Query Match     63.5%;  Score 1758.2;  DB 6;  Length 7788;
Best Local Similarity   87.9%;
Matches 1997;  Conservative    0;  Mismatches 168;  Indels 108;  Gaps    3;
```

```
Qy          6 CGCTGTGGAATCGTGTGTCAGTTAGGGTGTGGAAAGTCCCAGGCTCCCCAGGCAGGCAGAA 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 CGCTGTGGAATCGTGTGTCAGTTAGGGTGTGGAAAGTCCCAGGCTCCCCAGGCAGGCAGAA 72

Qy         66 GTATGCAAAGCATGCATTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 ATCCAGATCCACAAACCTTGCTTCAAAAAATGAACAACTTTACCGACGCGCGCCGGTTT 1205
```

```
Db   1045  ATCCAGATCCACAACCTTCGCTTCAAAAAATGAACACTTACCGACGCGCCGGTTT   1104
Qy   1206  ATCATCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTGTAGTCTCAGTGAGCCATATCCTTG   1265
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105  ATCATCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTGTAGTCTCAGTGAGCCATATCCTTG   1164
Qy   1266  CCTGATACCTGGCGAGATGGAAACCTTCTTGGCAACCGCTTCCCCGACTTCCTTACGAGGGG   1325
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165  CCTGATACCTGGCGAGATGGAAACCTTCTTGGCAACCGCTTCCCCGACTTCCTTACGAGGGG   1224
Qy   1326  AAGGCCCACCAGAAGCAAATTTCGTGTGTAAATTAGATAAAATCGTATTTTGTCAAATCACAGAGTGCT   1385
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225  AGGCCCACCAGAAGCAAATTTCGTGTGTAAATTAGATAAAATCGTATTTTGTCAAATCACAGAGTGCT   1284
Qy   1386  TTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCCCGCACTTTGAAATCTTGTAATCCTTG   1445
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285  TTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCCCGCACTTTGAAATCTTGTAATCCTTG   1344
Qy   1446  AAGGCTCCTCAGAAACAAGCTCTCTTCTTCAAAATCTATACATTAAGACGACTCGAAATTCCACA   1505
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345  AAGGCTCCTCAGAAACAGCTCTCTTCTTCAAAATCTATACATTAAGACGACTCGAAATTCCACA   1404
Qy   1506  TATCAAAATATCCGAGTGTAGTAGTGTAAACATTCCAAAACCGTGATGGAAATGGAACACACTTAA   1565
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405  TATCAAAATATCCGAGTGTAGTAGTGTAAACATTCCAAAACCGTGATGGAAATGGAACACACTTAA   1464
Qy   1566  AATGCGAGTATCCGGAATGATTGATTTGGACGACCGATGGCATGCGGAGAGAATCT   1625
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465  AATTGCGAGTATCCGGAATGATTGATTTGGACGAACGGTGTACATTGAACGACTCGAAATTCGGAGAATCT   1524
Qy   1686  GTTCATGATCAGTGCAATTGTCTTGTCCCTATCGGAACGACTCTGGCACAAAAATCGTATTC   1745
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525  CACGGCAGGCAGTTCTTATGAGGCAGCAGCCGAGCAGCCACCTTTAGGCAGTAGATCCAGAGGA   1584
Qy   1746  ATTAAAAACCGGAGGTAGATGAGATGTGACGTAGTGCATCGAACGTGTACATCCCTGGTA   1805
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585  GTTCATGATCAGTGCAATTGTCTTGTCCCTATCGGAACGACTCTGGCACAAAAATCGTATTC   1644
Qy   1806  ATCCGGTTTAGAATCCATGATAATAATTTTTTGGATGAATGGAGAGCTTTTTTTGCACGTT   1865
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1605  ATTAAAAACCGGGAGGTAGATGAGATGTGACGTAGTGCATCGAACGTGTACATCCCTGGTA   1704
Qy   1866  CAAAAATTTTTTGCAACCCCTTTTTTGGAAAACGAACCACACGGTAGGCTCGGCGAAAATGCCCAT   1925
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1705  ATCCGGTTTAGAATCCATGATAATAATTTTTTGGATGAATGGAGAGCTTTTTTTGCACGTT   1764
Qy   1926  ACTGTTGAGGCAATCACGTTCAATTCATTATAAATGTCGTTCCGTTCCGGGCGGCCGCAACTCCGAT   1985
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1765  CAAAAATTTTTTGCAACCCCTTTTTTGGAAAACGAACCACACGGTAGGCTCGGCGAAAATGCCCAT   1824
Qy   1986  AAATAACGGCGCCAAACACGGCATAAAAGAATTGAAGAGAGTTTCACTGCCATACGGACGAT   2045
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1825  ACTGTTGAGGCAATTCACGTTCAATTCATTATAAATGTCGTTCCGTTCCGGGCGGCCGCAACTCCGAT   1884
Qy   2046  TCTGTGATTTGTTGTTATTCAGTTCATTTCTGCGTTGCCAACCGAACCGGACATTTC   2105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1885  AAATAACGGCGCCAAACACGGCATAAAAGAATTGAAGAGAGTTTCACTGCCATACGGACGAT   1944
Qy   2106  GAAGTACTCAGCGTAAGTGAATGAGTCCATCGGATAGTGCATCTGTTAAAGCAAATTGTTCC   2165
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1945  TCTGTGATTTGTTGTTATTCAGTTCATTTCTGCGTTGCCAACCGAACCGGACATTTC   2004
Qy   2166  AGGAACCAGGCGGTATCTCTTCATTAGCCTTATGCAGTTGCTCTTCCAGCGCGGTTCCATCTTC   2225
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2005  GAAGTACTCAGCGTAAGTGAATGAGTCCATCGGATAGTGCATCTGTTAAAGCAAATTGTTCC   2064
Qy   2226  CAGCGGGATAGAAATGGCGCGCGGGGCCTTTCTTTATGTTTTTTGGCGTCTTCCATGG   2278
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2065  AGGAACCAGGCGGTATCTCTTCATTAGCCTTATGCAGTTGCTCTTCCAGCGCGGTTCCATCTTC   2124
Db   2125  CAGCGGGATAGAAATGGCGCGCGGGGCCTTTCTTTATGTTTTTTGGCGTCTTCCATGG   2177
```

RESULT 10
AAD31132

Query Match   63.1%;   Score 1747.2;   DB 6;   Length 7312;
Best Local Similarity   87.8%;   Pred. No. 3.3e-225;
Matches 1997;   Conservative 0;   Mismatches 168;   Indels 109;   Gaps 4;

```
Qy    6  CGCGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCAGGGCTCCCAGGCCAGCAGAA   65
```

```
Qy  2225 CCAGCGCGGATAGAATGGCGCCGGCGGCCTTTGCTTTTATGTTTTGTTTTTGGGGTCTTCCATGG 2278
         |||||||||||||||||||||||| |||||| ||||| ||| ||||||||||||||||||||||||
Db  2125 CCAGCGGATAGAAATGGCGCCGGCGCCCTTTCTTTAPGTTTTGTTTTTGGGGTCTTCCATGG 2178
```

RESULT 11

ID ADJ25519/c   ADJ25519 standard; DNA; 11004 BP.
XX
AC ADJ25519;
XX
DT 20-MAY-2004 (first entry)
XX
DE ARE-LUC/CMV-rAR transgenic construct DNA sequence SEQ ID NO:1.
XX
KW transgenic non-human mammal; androgen response element ; ARE;
KW androgen receptor; transgenic mouse;
KW selective androgen receptor modulator; SARM; cancer;
KW defective androgen receptor function disorder; transgenic construct;
KW ARE-LUC/CMV-rAR; gene; ds.
XX
OS Synthetic.
XX
PN WO200407753-A2.
XX
PD 22-JAN-2004.
XX
PF 16-JUL-2003; 2003WO-US022142.
XX
PR 17-JUL-2002; 2002US-0396501P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Attar RM, Bol DK, Gottardis M, Mookhtiar K, Ostrowski J;
PI Rowley RB;
XX
DR WPI; 2004-122974/12.
XX
PS Example; SEQ ID NO 1; 49pp; English.
XX

CC The present invention describes a transgenic non-human mammal whose
CC genome comprises a nucleic acid construct comprising a reporter nucleic
CC acid encoding a reporter operably linked to a promoter comprising an
CC androgen response element (ARE), and where the construct further
CC comprises an androgen receptor nucleic acid encoding an androgen
CC receptor, where expression of the reporter nucleic acid is regulated by
CC expression of the androgen receptor nucleic acid. The mammal expresses
CC the receptor nucleic acid in organs when the androgen receptor nucleic
CC acid is expressed. Also described: (1) a cell isolated from the
CC transgenic mouse, where the genome of the cell comprises the nucleic acid
CC construct; (2) a mouse cell line comprising the cell of (1); (3) an
CC isolated nucleic acid construct that comprises a reporter nucleic acid
CC encoding a reporter operably linked to a promoter comprising an ARE; (4)
CC obtaining a reporter mouse whose genome comprises the nucleic acid
CC construct, where the mouse can be bred to produce progeny mice whose
CC genomes comprise the nucleic acid construct; (5) producing a transgenic
CC mouse cell line that expresses a reporter nucleic acid; and (6) screening
CC for a modulator of the androgen receptor. The transgenic non-human mammal
CC can be used as an in vivo model for the identification and development of
CC selective androgen receptor modulators for the treatment of cancer or
CC other disorders associated with defective androgen receptor functions.
CC Selective androgen receptor modulators (SARMs) can act as antagonists or
CC agonists in different tissues containing the androgen receptor. The
CC present sequence represents a transgenic construct designated ARE-LUC/CMV
CC -rAR, which is used in the exemplification of the present invention.
XX
SQ Sequence 11004 BP; 2793 A; 2573 C; 2621 G; 3017 T; 0 U; 0 Other;

Query Match   60.0%; Score 1663.4; DB 12; Length 11004;
Best Local Similarity   97.4%;
Matches 1691; Conservative   0; Mismatches 46; Indels   0; Gaps   0;
```

```
Qy   542 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAATGGCCTAAGAG  601
Db  4385 ATATAACATGAAATTTTACAATAGGGCAAAAAGGAAGGAACAATCAAGGGTGCCCAAAACTCAC  4326

Qy   602 GCCGGAGTGTTTTACCCCAACCTTTAAACCGGCGGAGCGATCTTTCGCGCCCTTTCTTGGGCCTTTATGG  661
Db  4325 CCTGAAGTTCTCAGCTCTAGAAATTACACGGCGGATCTTTCCGCGCCCTTTCGTGGCCTTTATGA  4266

Qy   662 GGATCTCTTCCTGATTTTTTTCTTGCCGTCGGGATTTTTTCCGGTAAAGGCTTTCGTGGTTACTTCGTCCA  721
Db  4265 GGATCTCTTCCTGATTTTTTTCTTGCCGTCGGGAGTTTTTTCCGGTAAGAGACTTTTTCGTTCGTCCA  4206

Qy   722 CAAACAACAACTCCTTCGGCGCAACTTTTTTGGCGTCGGAGTTTTTTCGGTTTCTTGGTTTCGTTAATCCA  781
Db  4205 CAAACAACAACTCCTTCCGGCGCAACTTTTTTCGGCGGTTGTTACTTGGACTGGGCGTAATCCA  4146

Qy   782 CGATCTCTTTTTTCCGTCATCGTCTTCCGTGCTCCAAAACAACAACAACGGCGGGGGGAAGTT  841
Db  4145 CGATCTCTTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACAACAACGGCGGGGGTGGGAAGTT  4086

Qy   842 CACCCGCGTCATCGTCGGGGAAGACCTGCGACCACCTGCGTCGGAAGATGTTGGGGGTTGGA  901
Db  4085 CACCCGCGTCATCGTCGGGGAAGACCTGCGACACCTGCGTCGGAAGATGTTGGGGGTTGGA  4026

Qy   902 GCAAGATGGGATTCCAATTCAGGCGGAGCCAACTGGGCGTTTGTTACTTGGTTTTCAGAGACT  961
Db  4025 GCAAGATGGATTCCAATTCAGCGCGGAGCCACCTGATAGCCTTTGTTACTTAATCAGAGACT  3966

Qy   962 TCAGCGCGGTCAACCGATGAAGAAGTTCGTCTTCGTCCAGTAAGCTAGTGTCTCCAGAAT 1021
Db  3965 TCAGCGCGGTCAACCGATGAAGAAGTGTTCGTCTTCCGTCCCAGTAAGCTAGTGTCTCCAGAAT  3906

Qy  1022 GTAGCGCCATCCATCCTTGTCAATCAAGGCCGTTGGGTGTTGCTTCCGGATTGTTTTACATAACCGG 1081
Db  4085 CACCCGCGTCATCGTCGGGGAAGACCTGCGACACCTGCGTCGGAAGATGTTGGGGGTTGGA  4026

Qy  1082 ACATAATCATTAGGACCTCTCACACACAGTTCGCCCTTCTTTGATTAACGCCCCAGCGTTTCC 1141
Db  4025 GCAAGATGGATTCCAATTCAGCGCGGAGCCACCTGATAGCCTTTGTTACTTAATCAGAGACT  3966

Qy  1142 CGGTATCCAGATCCACAACCTTTCGCTTCAAAAAAATGGAACAAACTTACCGGACCGGCCCCG 1201
Db  3845 ACATAATCATTAGGACCTCTCACACACAGTTCGCCCTTCTTTGATTAACGCCCCAGCGTTTCC  3786

Qy  1202 GTTTATTCATCCCCTCGGGTTGTAATCAGAATAGTAGTGTTCAGTGAGCCCATATC 1261
Db  3785 CGGTATCCAGATCCACAACCTTTCGCTTCAAAAAAATGGAACAAACTTACCGGACCGGCCCCG  3726

Qy  1262 CTTGCCTGATACCTGGCAGATGGAACCTCTTGGCGAAGCCCGGACACTTCCTTAGAGA 1321
Db  3725 GTTTATTCATCCCCTCGGGTTGTAATCAGAATAGTAGTGTTCAGTGAGCCCATATC  3666

Qy  1322 GGGGAGGCGCCACCAGAAGCAAATTCGGTGTAAATTAGATAAAATCGTATTGTCAATCAGAG 1381
Db  3665 CTTGCCTGATACCTGGCAGATGGAACCTCTTGGCGAAGCCCGGACACTTCCTTAGAGA  3606

Qy  1382 TGCTTTTGGGGAAGAGCGGAAATAGGGTTGGCACCAGCAGCGCACTTGAAATCTTGTAAT 1441
Db  3605 GGGGAGGCGCCACCAGAAGCAAATTCGGTGTAAATTAGATAAAATCGTATTGTCAATCAGAG  3546

Qy  1442 CCTGAAGGCTCCTCAGAAACAGAGCTCTTCTTCAAATACTATAGGACGCTCGGAAATC 1501
Db  3545 TGCTTTTGGGGAAGAGCGGAAATAGGGTTGGCACCAGCAGCGCACTTGAAATCTTGTAAT  3486

Qy  1502 CACATATCAAAATATCCGAGTGTAGTGTAAACATTCAAAAACCGGTGATGGGAATGGAACAAACAC 1561
Db  3485 CCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTGAAGACGATCGAAATC  3426

Db  3425 CACATATCAAAATATCCGAGTGTAGTGTAAACATTCAAAAACCGGTGATGGGAATGGAACAAACAC  3366
```

RESULT 12

```
AC   AAD27536;
ID   AAD27536 standard; DNA; 4818 BP.
DT   18-APR-2002  (first entry)
DE   pGL3 basic vector DNA.
KW   firefly; ds.
KW   cell cycle control; DNA damage repair; pGL3 basic vector; apoptosis;
KW   p53 protein; pGL3 luciferase reporter vector; luc+; transcription factor;
OS   Chimeric.
OS   Unidentified.
OS   Photinus pyralis.
OS   Photinus pyralis.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..58
FT                   /*tag= a
FT                   /note= "Multiple cloning site"
FT   misc_feature    88..1737
FT                   /*tag= b
```

(rotated sequence alignment block — Qy/Db rows with sequence data)

```
Qy  1562 TTAAAAATGCAGTATCCGGAATAGTGTTGATTTGATTTGCAAAAATAGGAATTCTCGGGCATTGACANGCGGACA 1621
Db  3365 TTAAAAATGCAGTATCCGGAATAGTGTTGATTTGATTTGCAAAAAATAGGATTCTTGGCATGGCGGAGA 3306

Qy  1622 ATCTCACGCGAGGCAGTTCTATGAGGCAGCGAGCGACCAGCTTAGGCAGGACCAGTAGATCCAG 1681
Db  3305 ATCTCACGCGAGGCAGTTCTATGAGGCAGCGAGCGACACCTTTAGGCAGGACCAGTAGATCCAG 3246

Qy  1682 AGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCGT 1741
Db  3245 AGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCGT 3186

Qy  1742 ATTCATTAAAAACCCGGGAGGTAGATGTGAGGAGAGTGTGACTGTACATCGAAATCCCT 1801
Db  3185 ATTCATTAAAAACCCGGGAGGTAGATGTGACGAACGTGTACATCGAAATCCCT 3126

Qy  1802 GGTAATCCGTTTAGAATCCATGATAATAATAATTTTGGATGATGGGGAGCTTTTTGCA 1861
Db  3125 GGTAATCCGTTTAGAATCCATGATAATAATAATTTTGGATGATGGGGAGCTTTTTGCA 3066

Qy  1862 CGTTCAAAATTTTTTGCAACCCCTTTTTGGAAAACGAACACCACGTAGGCTGCGAAATGC 1921
Db  3065 CGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGTAGGCTGCGAAATGC 3006

Qy  1922 CCATACTGTTTGACAATTCACGTTCATTATAATAAATGTCGTTCGCGGCGAACTGCAACTC 1981
Db  3005 CCATACTGTTTGACGAATTCACGTTCATTATAAATGTCGTTCGCGGCGAACTGCAACTC 2946

Qy  1982 CGATAAAATAACGCGCCCAACACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACGA 2041
Db  2945 CGATAAAATAACGCGCCCAACACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACGA 2886

Qy  2042 CGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGACA 2101
Db  2885 CGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGACA 2826

Qy  2102 TTTCGGAACTACTCAGCGTAAGTGATGTCACCTCCGATAATGTGCATCTGTAAAAGCAATTG 2161
Db  2765 TTTCGGAACTACTCAGCGTAAGTGATGTCACCTCCGATAATGTGCATCTGTAAAAGCAATTG 2706

Qy  2162 TTCCAGGAACCAGGGCGTATCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCAT 2221
Db  2825 CGATAAAATAACGCGCCCAACACCCGGCCATAAAGAATTGAAGAGAGTTTTCACTGCATACGA 2766

Qy  2222 CTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATTTGTTTGGCGTCTTCCATGG 2278
Db  2705 CTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATTGTTTGGCGTCTTCCATGG 2649
```

```
FT   primer_bind
FT                   /note= "luciferase gene (luc+)"
FT                   complement(89..111)
FT                   /*tag= c
FT   primer_bind
FT                   /bound_moiety= "GL primer2"
FT                   1772..1993
FT                   /note= "SV40 late poly(A) signal"
FT                   /*tag= d
FT   polyA_signal
FT                   complement(2061..2080)
FT                   /*tag= e
FT   primer_bind     2318
FT                   /bound_moiety= "RV primer4"
FT                   /*tag= f
FT   misc_signal
FT                   complement(3083..3940)
FT                   /note= "ColE1-derived plasmid replication origin"
FT   misc_feature    4073..4527
FT                   /*tag= g
FT                   /note= "Beta-lactamase gene"
FT   misc_signal     4658..4811
FT                   /*tag= h
FT                   /note= "F1 origin"
FT   polyA_signal    4760..4779
FT                   /*tag= i
FT                   /note= "Upstream poly(A) signal"
FT   primer_bind
FT                   /*tag= j
FT                   /bound_moiety= "RV primer3"
XX
PN   WO200196602-A2.
XX
PD   20-DEC-2001.
XX
DR   WPI; 2002-130743/17.
XX
PI   Yang AL,  Festing M;
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PR   16-JUN-2000; 2000GB-00014820.
XX
PF   18-JUN-2001; 2001WO-GB002718.
XX
PD   20-DEC-2001.
XX
PS   Claim 8; Page 34-36; 53pp; English.
XX
CC   The patent discloses methods for determining the p53 status of a sample
CC   which comprises providing a sample containing a pGL3 luciferase reporter
CC   vector and determining whether p53 binds to the pGL3 vector. p53 is a
CC   transcription factor that regulates many genes including those associated
CC   with cell cycle control, apoptosis and DNA damage repair. pGL3 reporter
CC   vectors contain a modified firefly luciferase cDNA designated luc+. p53
CC   protein binds to pGL3-basic vector and causes luciferase expression. The
CC   method is useful for determining the p53 status of a sample. It is also
CC   useful for assaying for mimetics or antagonists of p53 and for assaying
CC   for presence of activated p53 and/or DNA damage. The invention also
CC   relates to a method of modifying pGL3 vector which involves deletion or
CC   alteration of a CCCGGG motif of the pGL3 vector and/or deleting or
CC   altering a sequence within 20 bp sequence 5' or 3' of CCCGGG motif.
CC   nucleic acid having a sequence incorporating the CCCGGG motif is useful
CC   for conferring promoter activity or p53-responsiveness on a nucleic acid
CC   encoding a polypeptide of interest or in assays for p53 transcriptional
CC   activity. The present DNA sequence is pGL3 basic vector sequence
XX
SQ   Sequence 4818 BP; 1251 A; 1146 C; 1159 G; 1262 T; 0 U; 0 Other;
```

```
Qy   591 TGGCCTAAGAGGGCCGGAGTGTTTACCCCAACCTTTAAACGGCGATCTTTCCGCCCTTCTT 650
```

Query Match   59.8%; Score 1657.6; DB 6; Length 4818;
Best Local Similarity 98.9%;
Matches 1669; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Db  1773 TGCTCGGAAGGGGCGCGCCGGGCCCCCGACTCTAGAGATTACACGGGACGATTTCCGCCCTTCTT 1714

Qy   651 GGCCTTTATGAGGAGTCTCTCTGATTTTTTCTTGCGTTCGGTTTCCGGTAAGAAGACCTTTCGG 710
          ||||||||||||||||||||||||
Db  1713 GGCCCTTTATGAGGAGATCTCTCTGATTTTTTTTCTTGCGTTTTCTTGCGTCGGAGTTTTTCGG 1654

Qy   711 TACTTCGTCCACAAACACAACTCCTCGATTTTTTCTTGCGGTTGGTGTTGTTACTTGACTGGC 770
          ||||||||||||||
Db  1653 TACTTCGTCCACAAACACAACTCCTCCGCGGCAACTCTTTTTCTTGCGGTTGGTGTTACTTGGC 1594

Qy   771 GACGTAATCCACGATCTCTTTTTCCGTCGCTCCCAAAACAACAACGGC 830
          ||||||||||||
Db  1593 GACGTAATCCACGATCTCTTTTTTCCGTCGCTCCAAAAACAACAACGGC 1534

Qy   831 GGCGGGGAAGTTCACCGGGGTCATCGTCGTCGGGAAGAAGACCTGGCACACTTGTT 890
          |||||||||||||
Db  1533 GGCGGGGAAGTTCACCGGGGTCATCGTCGGGGAAGAAGACCTGGCACACTTGTT 1474

Qy   891 GGGGTGCTTGGAGCAAGATGGATTCCAATTCAGCCGGCTTTTTGTACTT 950
          |||||||||||||||
Db  1473 GGGGGTGTTGGAGCAAGATGGATTCCAATTCAGCCGGCTTTTTGTACTT 1414

Qy   951 AATCAGAGACTTCAGGCGGTCAACGATGAAGAAGAGTTGTTCGTCTTGTTGTTAGCTAT 1010
          |||||||||||||||
Db  1413 AATCAGAGACTTCAGGCGGTCAACGATGAAGAAGAGTTGTTCGTCTTGTTAGCTAT 1354

Qy  1011 GTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCCTTGGGTGCGCTTCCGGATTGTT 1070
          ||||||||||||
Db  1353 GTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCCTTGGGTGCGCTTCCGGATTGTT 1294

Qy  1071 TACAATAACCGGACATAATCATAGGAGACCTTCTCACACAGTTCGCCTTCACACAGTTAACGCC 1130
          ||||||||||
Db  1293 TACAATAACCGGACATAATCATAGGAGACCTTCTCACACAGTTCGCCTTTGATTAAACGCC 1234

Qy  1131 CAGCGTTTTCCCGGTATCCAGATCCACAACTTCGCTTCAAAAAATGGAACAACTTTACC 1190
          ||||||||||
Db  1233 CAGCGTTTTCCCGGTATCCAGATCCACAACTTCGCTTCAAAAAATGGAACAACTTTACC 1174

Qy  1191 GACCGCGCCGGGTTTATCATCCCCCTCGGGTGTAATCAGAATAGTCGATGTTAGTCTCAGT 1250
          ||||||||||
Db  1173 GACCGCGCCGGGTTTATCATCCCCCTCGGGTGTAATCAGAATAGTCGATGTTAGTCTCAGT 1114

Qy  1251 GAGCCCATATCCTTGCCTGATACTGGGCGGAACCTGCTTCTTGCGCTTGCGTTCCCGAC 1310
          ||||||||||
Db  1113 GAGCCCATATCCTTGCCTGATACTGGGCGGAACCTGCTTCTTGCGCTTGCGTTCCCGAC 1054

Qy  1311 TTCCTTAGAGAGGGGGGGCGGCCCCAGAAGCCAATTTCGTGTAAATTAGATAAATCGTTATTT 1370
          ||||||||||
Db  1053 TTCCTTAGAGAGGGGGGGCGGCCCCAGAAGCCAATTTCGTGTAAATTAGATAAATCGTTATTT 994

Qy  1371 GTCAAATCAGAGTGCTTTTGGCGAAGAAGGAGGAAGAATAGGGGTTGGCACCGGACACTTTG 1430
          ||||||||||
Db   993 GTCAAATCAGAGTGCTTTTGGCGAAGAAGGAGGAAGAATAGGGGTTGGCACCGGACACTTTG 934

Qy  1431 AATCTTGTAATCCTGAAGGCTCCTCCAGAAAACAGCTCTTCCTTCAAATCTATACATTAAGAC 1490
          ||||||||||
Db   933 AATCTTGTAATCCTGAAGGCTCCTCCAGAAAACAGCTCTTCCTTCAAATCTATACATTAAGAC 874

Qy  1491 GACTCGGAAATCCACATAATCAAATATTCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAA 1550
          ||||||||||
Db   873 GACTCGGAAATCCACATAATCAAATATTCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAA 814

Qy  1551 TGGAACAACACTTAAAATGCGGAATCGTATCCGGAATGAATTGCAAAAATAGGATCTCT 1610
          ||||||||||
Db   813 TGGAACAACACTTAAAATGCGGAATCGTATCCGGAATGAATTGCAAAAATAGGATCTCT 754

Qy  1611 GGCATGCGAGAATCTCACGCAGGCAGTTCTATGAGGCAGCACACCTTTAGGCAGACC 1670
          ||||||||||
Db   753 GGCATGCGAGAATCTCACGCAGGCAGTTCTATGAGGCAGCACACCTTTAGGCAGACC 694

Qy  1671 AGTAGATCCAGAGGAGAGTTCATTGATCAGGTCAATTGTCTTCCTATCGAAGGACTCTG 1730
          ||||||||||
Db   693 AGTAGATCCAGAGGAGAGTTCATTGATCAGGTCAATTGTCTTCCTATCGAAGGACTCTG 634
```

```
Qy  1731 CACAAAATCGTATTCATTAAAAACCCGGCAGCGGTGAGTGAGTGAGTGAGTGTACATCGA 1790
         ||||||||||
Db   633 CACAAAATCGTATTCATTAAAAACCCGGCAGCGGTGAGTGAGTGAGTGAGTGTACATCGA 574

Qy  1791 CTGAAAATCCCTTGGTAATCCGTTTTTAGAAATCATTGATAATAATTTTTTGGGGAG 1850
         ||||||||||
Db   573 CTGAAAATCCCTTGGTAATCCGTTTTTAGAAATCATTGATAATAATTTTTTGGGAG 514

Qy  1851 CTTTTTTCGGCAGTTCAAAATTTTTGCAACCCCTTTTTTGGAAACGAAACACCACGGTAGG 1910
         ||||||||||
Db   513 CTTTTTTCGGCAGTTCAAAATTTTTGCAACCCCTTTTTTGGAAACGAAACACCACGGTAGG 454

Qy  1911 CTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTACAGTTCGTTCGTTCGCGCGCGC 1970
         ||||||||||
Db   453 CTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTACAGTTCGTTCGTTCGCGCGC 394

Qy  1971 AACTGCAACTCCGATAAAATAACGCGCCCACACACCGCCATAAAAGAATTGAAGAGTTTTC 2030
         ||||||||||
Db   393 AACTGCAACTCCGATAAAATAACGCGCCCACACACCGCCATAAAAGAATTGAAGAGTTTC 334

Qy  2031 ACTGCATACGACGATTCTGTGATTTGTATTCAGCCCATATGTCTTTATTAGTTCTGCCAA 2090
         ||||||||||
Db   333 ACTGCATACGACGATTCTGTGATTTGTATTCAGCCCATATGTCTTTATTAGTTCATAGCCAA 274

Qy  2091 CCGAACGGACATTTCGAAGTACTCACGGTAAGTGAGTGATGTCCACCTCGATATGTGCATCTGT 2150
         ||||||||||
Db   273 CCGAACGGACATTTCGAAGTACTCACGGTAAGTGAGTGATGTCCACCTCGATATGTGCATCTGT 214

Qy  2151 AAAAGCAAATTGTTTCCAGGAACCAGGCGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCC 2210
         ||||||||||
Db   213 AAAAGCAAATTGTTTCCAGGAACCAGGCGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCC 154

Qy  2211 AGCGGTTCCATCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTTTTCAAGGTC 2270
         ||||||||||
Db   153 AGCGGTTCCATCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTTTTGTGTCC 94

Qy  2271 TTCCATGG 2278
         ||||||||
Db    93 TTCCATGG 86
```

```
PT   heat responsive element, Myc, p53 or alkaline phosphatase 1.
XX
PS   Claim 7; Page 39-42; 42pp; English.
XX
CC   A new construct for measuring transcription comprises a cis-element from
CC   a DNA sequence such as that for nuclear factor-kappaB (NF-kappaB), heat
CC   responsive element (HRE), Myc, p53, nuclear factor of activated T-cells
CC   (NFAT), activator protein 1 (AP-1), serum responsive element (SRE),
CC   glucocorticoid responsive element (GRE) or cyclic adenosine monophosphate
CC   (cAMP) responsive element, and a reporter gene and a promoter. Reporter
CC   genes are chosen from secreted alkaline phosphatase (SEAP), destabilized
CC   green fluorescent protein (d2EGFP) and luciferase (luc). Promoter are
CC   preferably chosen from the thymidine kinase and gonadotropin-alpha gene
CC   promoters. The constructs are useful for measuring transcription and for
CC   monitoring the activation of a transcription factor. They are also used
CC   to determine whether a gene is involved in signal transduction pathway,
CC   where the cis-element is a component of a DNA sequence which affects the
CC   expression of a test gene
XX
SQ   Sequence 4987 BP; 1295 A; 1188 C; 1207 G; 1297 T; 0 U; 0 Other;

Query Match    59.8%;  Score 1657.6;  DB 3;  Length 4987;
Best Local Similarity  98.9%;  Pred. No. 3.3e-213;
Matches 1669;  Conservative    0;  Mismatches 19;  Indels    0;  Gaps    0;

Qy      591 TGGCCTAAGAGAGGGCCGGAGTGTTTACCCGAACCTTAAACGGGCGGATCCAATTCTTCGCCCTTCTT 650
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1942 TGCTCGAAGCGGGGCGCCCGCCCCCGGACTCTAGAATTACACGGCGGGATCTTTCGGGCGCCCTTCTT 1883

Qy      651 GGCCTTTATGAGGGATCTCTCTGATTTTTCTTGCCTCGAGTTTTTCGGGTAAGACCTTTCGG 710
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1882 GGCCTTTATGAGGGATCTCTCGATTTTTCTTGGGTCGAGTTTTTCGGGTAAGACCTTTCGG 1823

Qy      711 TACTTCGTCCACAAAACAAACTCCTCCGCGGCCAACTCTTTTTTCGGCGGTTGTTACTTGACTCGCC 770
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1822 TACTTCGTCCACAAAACAAACTCCTCCGCGGCCAACTCTTTTTTCGCGGTTGTTACTTGACTCGCC 1763

Qy      771 GAGGTAATCCACGATCTCTTTTTTCCGTCGGGGTTCCGTGCTCATCGTGCGTCTCAAAACAACGGC 830
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1762 GAGGTAATCCACGATCTCTTTTTTCCGTCGGGGTTCCGTGCTCATCGTGCGTCTCAAAACAACGGC 1703

Qy      831 GGGCGGGAGAGTTCACCGGCGCTCATCGTCGGGACGACCTGGGACGACCTGCTCGAAGATGTT 890
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1702 GGGCGGGAGAGTTCACCGGCGCTCATCGTCGGGAAGACCTGGGACCTGCTCGAAGATGTT 1643

Qy      891 GGGGTGTTGGGAGCAAGATGGATTCCAAATTCAGCGGGGAGCCACCTGGCTGAGCCTTTGTACTT 950
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1642 GGGGTGTTGGGAGCAAGATGGATTCCAAATTCAGCGGGGAGCCACCTGATAGCCTTTGTACTT 1583

Qy      951 AATCAGAGACTTCAGGCGGGTCAACGGATGAAGAAGTGTTCGTGTTCTTCGTCCCAGTAAGCTAT 1010
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1582 AATCAGAGACTTCAGGCGGGTCAACGGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTAT 1523

Qy     1011 GTCTCCAGNATGTTAGCCATCCATCCTTCGTCGTTTGTCGCTTCGCTTCCGGATTGTT 1070
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1522 GTCTCCAGANATGTAGCCATCCATCCTTGTCAATCAAGGGGTTGTTGGTCGCTTCCGGATTGTT 1463

Qy     1071 TACATAACCGGACACATAATCATAGGACCTCACACACAGTGCGCTCTTTGATTAACGCC 1130
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1462 TACATAACCGGACATAATCATAGGACCTCACACACAGTGCGCTCTTTGATTAACGCC 1403

Qy     1131 CAGGGTTTCCCGGTATCCAGATCCACAACCTTCCTTCAAAAAAATGGAACACTTTACC 1190
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1402 CAGGGTTTTCCCGGTATCCAGATCCAGAGTACTCAGCGTAAGTGAAGGTAGTCGCATCGA 1343

Qy     1191 GACCGGCGCCCGGTTTATCATCCCCTCGGGTGTTAATCAGATAACTGAGTCTCAGT 1250
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1342 GACCGGCGCCCGGTTTATCATCCCCTCGGGGTGTTATCAGATAACTGAGTAGTCTCAGT 1283

Qy     1251 GAGCCCATATCCTTGCCTGATACCTGGCAGAATGGAACCTCTTGGCCAACCGCTTCCCCGAC 1310
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1282 GAGCCCATATCCTTGCCTGATACCTGGCAGAATGGAACCTCTTGGCCAACCGCTTCCCCGAC 1223

Qy     1311 TTCCTTAGAGAGGGGAGGGCGCCACCAGAGAGCAATTTCGTGTAAAATTAGATAAAATCGTATTT 1370
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1222 TTCCTTAGAGAGGGGAGGGCGCCACCAGAGAGCAATTTCGTGTAAAATTAGATAAAATCGTATTT 1163

Qy     1371 GTCAATCAGAGTGCTTTTGGCGAAGAAGGAAGAATAGGGGTTGGGCACCAGCAGCCACTTTG 1430
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1162 GTCAATCAGAGTGCTTTTGGCGAAGAAGGAAGAATAGGGGTTGGGCACCAGCAGCCACTTTG 1103

Qy     1431 AATCTTGTAATCCTCGAAGGCTCCTCAGAGAAACAGCTCTTCTTCAAATCTATACATTAAGAC 1490
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1102 AATCTTGTAATCCTCGAAGGCTCCTCAGAGAAACAGCTCTTCTTCAAATCTATACATTAAGAC 1043

Qy     1491 GACTCGAAATTCCACATATCAAAATATCCCAGTGTAGTGAAACATTCCAAAAACCGTGATGGAA 1550
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1042 GACTCGAAATTCCACATATCAAAATATCCCAGTGTAGTGAAACATTCCAAAAACCGTGATGGAA 983

Qy     1551 TGGAACAACACTTAAAATCGGAAATGAATGGACTATCCGGAATGATTTGGCCAAAAATAGGATCTCTT 1610
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      982 TGGAACAACACTTAAAATCGGAAATGAATGGACTATCCGGAATGATTTGGCCAAAAATAGGATCTCT 923

Qy     1611 GGCATGCGAGAAATCTCACCGCAGCCAGTTCTTATGAGGCAGAGCGACACCTTTAGGCAGACC 1670
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      922 GGCATGCGAGAAATCTCACCGCAGCCAGTTCTTATGAGGCAGAGCGACACCTTTAGGCAGACC 863

Qy     1671 AGTAGAATCCAGAGAGAGTTCATGATCATGATCAGTGCAATGTCTTGTGTCCCTATCGAAGGACTCTGG 1730
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      862 AGTAGATCCAGAGAGAGTTCATGATCAGTGCAATGTCTTGTGTCCCTATCGAAGGACTCTGG 803

Qy     1731 CACAAAATCGTATTCATTAAAACCCGGGAGGGTAGGTAAGTGAATGTGGTAAACATTCCA 1790
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      802 CACAAAATCGTATTCATTAAAACCCGGGAGGGTAGGTAAGTGAATGTGGTGTACATCGA 743

Qy     1791 CTGAAATCCCTGGTAATCCGTTTTGAATCCATGAATTAAATTTGGATGATTGATTGGGAG 1850
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      742 CTGAAATCCCTGGTAATCCGTTTTTGAGATCCATGAATTAAATTTGGATGATTGATTGGGAG 683

Qy     1851 CTTTTTTCGCACGTTCAAAATTTTTTGGAAACGAACACCACCAGGTAGG 1910
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db      682 CTTTTTTTTGCACGTTCAAAATTTTTTTGGAAACGAACACCACCAGGTAGG 623

Qy     1911 CTGCGAAATGCCCATACTGTTGAGCAATTCAGTTCATTGAAGATTAATAATTTTTGGGATTGATTGGGAG 1970
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      622 CTGCGAAATGCCCATACTGTTGAGCAATTCAGTTCATTAAAATGTCGTTCGGGCCGC 563

Qy     1971 AACTCCAACTCCGATAAAATAACGCGCCCCAAACACCGGCATAAAGAATTGAAGAGAGTTTTC 2030
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      562 AACTGCAAATCCGATAAAATAACGCGCCCCAAACACCGGCATAAAGAATTGAAGAGAGTTTTC 503

Qy     2031 ACTGCCATAGGACGATTCTCTGATTTGTATTCAGCGCCCATAATCGTTTCATAGGCTTCTGCCAA 2090
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      502 ACTGCATACGACGATTCTCTGATTTGTATTCAGCCCATAATCGTTTCATAGGCTTCTGCCAA 443

Qy     2091 CCGAACGGACATTCGAAGTACTCAGCGTAAGTAAGTGCCACTCGATATGTGCATCTGT 2150
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      442 CCGAACGGACATTCGAAGTACTCAGCGTAAGTAAGTGCCACTCGATATGTGCATCTGT 383

Qy     2151 AAAAGCAAATTGTTCCAGGAAACCAGGGCGTATCTCTTCATGCAGTTGCTCTCC 2210
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      382 AAAAGCAAATTGTTCCAGGAAACCAGGGCGTATCTCTTCATGCAGTTGCTCTCC 323

Qy     2211 AGCCGGTTCCATCTTCCAGCGGATAGAAATGGCGCGGGCCCTTTCTTATGTTTTTTGGGCGTC 2270
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      322 AGCCGGTTCCATCTTCCAGCGGATAGAAATGGCGCGGGCCCTTTCTTATGTTTTTTGGGCGTC 263

Qy     2271 TTCCATGG 2278
            ||||||||
Db      262 TTCCATGG 255
```

```
AC    ABV73856;
XX
DT    29-AUG-2003  (revised)
DT    08-JAN-2003  (first entry)
XX
DE    Plasmid pNFkB-Luc.
XX
KW    Screening; cell-based assay; cellular target; firefly; luciferase;
KW    reporter; pNFKB-Luc; gene; ds.
XX
OS    Photinus pyralis.
OS    Herpes simplex virus.
OS    Chimeric.
XX
XX
PN    WO2002272783-A2.
XX
PD    19-SEP-2002.
XX
PF    12-MAR-2002; 2002WO-US007713.
XX
PR    12-MAR-2001; 2001US-0275266P.
XX
PA    (IRMI-) IRM LLC.
XX
PI    Caldwell JS,  Chanda SK,  Somia NV,  Hogenesch JB,  Cooke MP;
PI    Aza-Blanc P;
XX
DR    WPI; 2002-750492/81.
XX
PT    Identifying targets of effectors of gene expression or cellular activity,
PT    by contacting reporter cells with an effector, adding nucleic acid
PT    encoding a target and identifying cells with altered expression or
PT    activity.
XX
PS    Example 1; Page 130-132; 140pp; English.
XX
CC    The present sequence is that of pNFkB-Luc, a plasmid designed for
CC    monitoring the activation of the NFKB signal transduction pathway. The
CC    plasmid contains the firefly luciferase (luc) gene (E354K mutant), and 4
CC    tandem copies of the NFKB consensus sequence fused to a TAT-like promoter
CC    from the herpes simplex virus thymidine kinase gene. A 1912 bp segment of
CC    the plasmid containing the NFKB consensus sequences, the promoter and the
CC    luciferase coding sequence was amplified by PCR (see ABV73854-55) and
CC    cloned into the XhoI and BamHI sites of the SIN retroviral reporter
CC    vector, producing SKBL-N. The vector was used for stable reporter cell
CC    generation in an example of the invention. The invention provides cell-
CC    based screening methods for identifying the targets of effectors or
CC    modulators of gene expression or cellular activity. The effect of a small
CC    molecule or other signal on a reporter cell is titrated by expressing
CC    within the cell cDNA that encodes a polypeptide that is the molecular
CC    target or that is responsible for directly or indirectly producing the
CC    molecular target. The method permits identification of gene function by
CC    modulating the level of message. NFKB reporter cells in a multiwell plate
CC    can be used. The method is fully automated and provides an increased
CC    throughput over conventional methods. Miniaturization and automation of
CC    transfection/transduction procedures permit comprehensive studies of
CC    phenotypes or pathways at the level of genome. (Updated on 29-AUG-2003 to
CC    standardise OS field)
XX
SQ    Sequence 4987 BP; 1295 A; 1188 C; 1207 G; 1297 T; 0 U; 0 Other;

Query Match             59.8%;  Score 1657.6;  DB 6;  Length 4987;
Best Local Similarity   98.9%;
Matches 1669;  Conservative  0;  Mismatches  19;  Indels  0;  Gaps  0;

Qy        591 TGGGCCTAAGAGGCCCGGAGTGTTTTAAACGGCGATCGTTTTCCGGATTCTTCTT 650
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1942 TGCTCGGAAGCGGCGCGGCGCCCCCGACTCTTAGAATTACAGGGCGGATCGTTTTTCCGGATTCTTCTT 1883
```

... (sequence alignment continues)

```
Qy        711 TACTTCGGTCCACAAACACAACAACTCCTCGCGGCGGCGAACCTTTTTTGTTTACTTGACTGGC 770
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1822 TACTTCGGTCCACAAAACACAACTCCTCGCGGCGGCGAACCTTTTTTGTTTACTTGACTGGC 1763

Qy        771 GACGTAATCCACGATCTCTTTTTCCGTCATCGTGCTCCAAAACGAACAACAACGGC 830
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1762 GACGTAATCCACGATCTCTTTTTCCGTCATCGTGCTCCAAAACGAACAACAACGGC 1703

Qy        831 GGCGGGAGAAGTTCACCGGCGTCATCGTCGGGAAGAGACCTGCCACACTCGGTGGGAAGAGTGTT 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1702 GGCGGGGAGAAGTTCACCGGCGTCATCGTCGGGAAGAGACCTGCCACACTCGGTGGGAAGAGTGTT 1643

Qy        891 GGGGTGTTGGAGCAAGATGGATTCCAAATTCAGCGCGGGAAGCGCACCCAACCTCAACTTGTACTT 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1642 GGGGTTGTTGGAGCAAGATGGATTCCAAATTCAGCGCGGGAAGCGCACCCAACCTCAACTTGTACTT 1583

Qy        951 AATCAGGAGACTTCAGCGCGGGTCGAACGATGAAGAAAGTGTTCGTCTTGATTAACGGC 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1582 AATCAGGAGACTTCAGCGCGGGTCAAACGATGAAGAAAGTGTTCGTCTTGATTAAGCTAT 1523

Qy       1011 GTCTTCCAGAATGTAGCGCCATCCATCCTTGTCAAATCAAGGCGGTTGGTCGCTTTGGTT 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1522 GTCTTCCAGAATGTAGCGCCATCCATCCTTGTCAAATCAAGGCGGTTGGTCGCTTTGGTT 1463

Qy       1071 TACATAACCCGGACATAATCAATCCATTAGGACCTTCTCACACACAGTTTCCGCTTCGGCTTT 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1462 TACATAACCCGGACATAATCAATCCATTAGGACCTTCTCACACACAGTTTCCGCTTCGGCTTT 1403

Qy       1131 CAGGCGTTTTCCCGGTATCCACGATCCACACCTTCGGCTTCAAAAAATGGAACAACTTTACC 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1402 CAGGCGTTTTCCCGGTATCCACGATCCACACCTTCGGCTTCAAAAAATGGAACAACTTTACC 1343

Qy       1191 GACCGCGCGCCCGGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGTAGTAGTCTCAGT 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1342 GACCGCGCGCCCGGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGTAGTAGTCTCAGT 1283

Qy       1251 GAGCCCCATATCCTTCGCCTGATACCTGGCAGAACCTCTTGCCAACCGCGTTCCCCGAC 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1282 GAGCCCCATATCCTTCGCCTGATACCTGGCAGAACCTCTTGGCAACCGCGTTCCCCGAC 1223

Qy       1311 TTCCTTAGAGAGGGGGAGCGCGCCACCAGAAGCAATTTCGTGTAAATTAGGACAACTGTATTT 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1222 TTCCTTAGAGAGGGGGAGCGCGCCACCAGAAGCAATTTCGTGTAAATTAGATAAAATCGTATTT 1163

Qy       1371 GTCAATCAGAGTGCTTTTGGCGAAGAAGAAGGAGAAATAGGGTTGGCACCAGCACTTTG 1430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1162 GTCAATCAGAGTGCTTTTGGCGAAGAAGAAGGAGAAATAGGGTTGGCACCAGCACTTTG 1103

Qy       1431 AATCTTGTAATCCTGAAGGCTCTCAGAAACAGCTCTTCTCAAATCTTATACATTAAGAC 1490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1102 AATCTTGTAATCCTGAAGGCTCTCAGAAACAGCTCTTCTCAAATCTATACATTAAGAC 1043

Qy       1491 GACTCGGAAATCCACATATCAAATATCGAGTGTAGTAGTCAAAAACCGTGATGGAA 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1042 GACTCGAAATCCACATATCAAATATCCGAGTGTAGTAGTCAAAAACCGTGATGGAA 983

Qy       1551 TGGAACAAACACTTAAAATCGGCAGTATCCGGAATGATTTGATTGCCAAAATAGGATCTCT 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 TGGAACAAACACTTAAAATCGGCAGTATCCGGAATGATTTGATTGCCAAAATAGGATCTCT 923

Qy       1611 GGCATGCGACGAATCTCACGCAGGCAGTTCTATGAGGCAGACGACCCTTTAGGCAGACC 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        922 GGCATGCGAGAATCTCACGCAGGCAGTTCTATGAGGCAGACGACCCTTTAGGCAGACC 863

Qy       1671 AGTAGATCCAGGAGAGTTCATGAGTCATGCAATTGCTTGTCCTGTCTGGAAGGACTCTGG 1730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        862 AGTAGAATCCAGAGGAGTTCATGATCAGTGCAATTGTCTTGTCCTATCGAAGGACTCTGG 803

Qy       1731 CACAAAAATCGTATTCATTAAAACCCGGAGGTAGATGAGTGTGAAGGTTACATCGA 1790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        802 CACAAAATCGTATTCATTAAAACCCGGAGGTAGATGAGATGAGAACCGTGTACATCGA 743

Qy       1791 CTGAAATCCCTGGTAATCCGTTTGAATCCATGATAATAATTTTTGGATGATTGGGGAG 1850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    742  CTGAAAATCCCTGGTAAATCCGTTTTTAGAARTCCATGATAATAATTTTTTGGATGATTGGGGAG  683
            |||||
Qy   1851  CTTTTTTTTTGCAGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGGAACACACCACGGTAGG  1910

Db    682  CTTTTTTTTGCACGTTCAAAATTTTTTTGCAACCCCTTTTTGGAAACGGAACACACGGTAGG    623
            |||||
Qy   1911  CTGGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTATATAAATGTCGTTCGCGGCGGC  1970

Db    622  CTGGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTATATAAATGTCGTTCGCGGCGGC   563
            |||||
Qy   1971  AACTGGCAACTCCGATAAATAAAATAACCGGCCCAACACCCGGCCCA                   2030

Db    562  AACTGCAACTCCGATAAATAAAATAACCGGCCCAACACCCGGCCCA                     503
            |||||
SQ Sequence 4987 BP; 1295 A; 1188 C; 1207 G; 1297 T; 0 U; 0 Other;

Qy   2031  ACTGCATACGACGATTCTGTGATTGTATTCAGCCCCATATCGTTTCATTAGCTTCTGCCAA     2090
            |||||
Db    502  ACTGCATACGACGATTCTGTGATTGTAATTCAGCCCCATATCGTTTCATTAGCTTCTGCCAA   443

Qy   2091  CCGAACGGACACATTTCGAAGTACTTCAGGCGGTAAGTAGTGATGTCCACCTTGCGATATGTGCATCTGT  2150
            |||||
Db    442  CCGAACGGACACATTTCGAAGTACTTCAGGCGGTAAGTGATGTCCACCTCTTAGAATTACACGGCGATCTTTTCGCGCCCTTTCTT  383

Qy   2151  AAAAGCAAATTGTTCCAGGAAACACAGGCGGGCGTAATCTCTTCATTAGCCTTATCGCAGTTGATTTTTTGCTCTCC  2210
            |||||
Db    382  AAAAGCAAATTGTTCCAGGAAACACAGGCGGGCGTAATCTCTTCATTAGCCTTATCGCAGTTGATTTTTTGCTCTCC  323

Qy   2211  AGGGGTTCATCTTCAGCGGGATAGAAATAGGGGGCCGCGGGCCGCTTTCTTANGTTTTTTTGGGGTC        2270
            |||||
Db    322  AGCGGTTCCATCTTCACGGGATAGAAATGGCGCCGGGGCCTTTCTTTATTGCTTTGGTTATTGACTGGC   263

Qy   2271  TTCCATGG  2278
            |||||
Db    262  TTCCATGG  255
```

RESULT 15
ABS76596/c

```
ID   ABS76596 standard; DNA; 4987 BP.
XX
AC   ABS76596;
XX
DT   11-DEC-2002   (first entry)
XX
DE   Human inducibly regulated gene promoter associated vector #1.
XX
KW   Responder cell; expression construct; screening assay; gene regulation;
KW   biosensor; reporter gene construct; vector; ds; circular; cyclic.
XX
OS   Synthetic.
XX
PN   WO20272789-A2.
XX
PD   19-SEP-2002.
XX
PF   12-MAR-2002; 2002WO-US008008.
XX
PR   12-MAR-2001; 2001US-0274979P.
PR   12-MAR-2001; 2001US-0275070P.
PR   12-MAR-2001; 2001US-0275148P.
XX
PA   (IRMI-) IRM LLC.
XX
PI   Caldwell JS, Su AI, Hogenesch JB;
XX
DR   WPI; 2002-723339/78.
XX
PT   Producing a collection of responder cells for high throughput screening
PT   assays, comprises identifying and cloning regulatory regions into
PT   expression constructs to control nucleic acids, and introducing into
PT   addressable cells.
XX
PS   Disclosure; Page 177-178; 187pp; English.
```

```
CC   The invention describes producing a collection of responder cells for
CC   high throughput screening assays, by identifying and cloning regulatory
CC   regions into expression constructs to control nucleic acids, and
CC   introducing the constructs into addressable cells. The method is useful
CC   in producing cells used in high throughput screening assays for profiling
CC   substances and conditions and for studying the function of the regulatory
CC   region mediating the response. The cells serve as biosensors to assess
CC   the effects of any perturbation, such as external or internal condition,
CC   on the cells from which the regulatory regions in the reporter gene
CC   constructs are derived can be inferred. This sequence represents a vector
CC   useful for the delivery of reporter gene constructs into cells
XX
SQ   Sequence 4987 BP; 1295 A; 1188 C; 1207 G; 1297 T; 0 U; 0 Other;
```

Query Match 59.8%; Score 1657.6; DB 6; Length 4987;
Best Local Similarity 98.9%; Pred. No. 3.3e-213;
Matches 1669; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy    591  TGGGCCTAAGAGGCCGGAGTGTTTACCCAACCTTTAAACGGGCGGATCTTTTCCGCGCCCTTCTT   650
            ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1942  TGCTCGGAAGCGGGCGGGCCGGGCGCCCCCGGACTCTTAGAATTACACGGCGATCTTTTCCGCGCCCTTTCTT  1883

Qy    651  GGGCCTTTATGAGAGATCTCTCTGATTTTTCTTGCCTGCGAGTTTTCCGGTAAGACCTTTTCGG   710
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1882  GGGCCTTTATGAGAGGATCTCTCTGATTTTTCTTGCGTCGAGTTTTCCGGTAAGACCTTTTCGG  1823

Qy    711  TACTTCGTCCACAAAACACAACAACTCCTCCGCGGCAACTTTTTCTGGGTGGTTGTTACTGGC   770
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1822  TACTTCGTCCACAAAACACAACTCCTCCGCGGCAACTTTTTCTGGGTGGTTGTTACTTGACTGGC  1763

Qy    771  GACGTTAATCCACGATCTCTTTTTCCGGTCATCGTGTCTTTCCGGTGCTCCAAAAACAACAACGGC   830
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1762  GACGTTAATCCACGATCTCTTTTTCCGGTCATCGTCTTTTCCGGTGCTCCAAAAACAACAACGGC  1703

Qy    831  GGCGGGAAGTTCACCGGCCTCATCGTTCGGAGAGACCTGCGACACCTGCGTCGCGAAGATGTT   890
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1702  GGCGGGGAAGTTCACCGGCGTCATCGTGCGGAAGAGACCTGCGACACCTGCGTCGCGAAGATGTT  1643

Qy    891  GGGGGTGTTGGAGCAAGATGGATTCCAATTCAGCGGGAGCGGGAGCCACCTGAGTAGCCTTTGTACTT   950
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1642  GGGGGTGTTGGAGCAAGATGGATTCCAATTCAGCGGGGAGCCACCTGAGTAGCCTTTGTACTT  1583

Qy    951  AATCAGAGACTTCAGGCGGTCAACGGATGAAGAAGTGTTCGGTTCGTCTTCGTCCCAGTAAGCTAT  1010
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1582  AATCAGAGACTTCAGGCGGTCAACGGATGAAGAAGTGTTCGGTTCGTCTTCGTCCCAGTAAGCTAT  1523

Qy   1011  GTCTCCAGAATGTGAGCCATCCATCCTTGTCAATCAAGGCGGTTGGTGTTCGCTTCCGGATTGTT  1070
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1522  GTCTCCAGAATGTGAGCCATCCATCCTTGTCAATCAAGGCGGTTGGTGTTCGCTTCCGGATTGTT  1463

Qy   1071  TACATAACCGGACATAATCATAGGACCTCTCACACACAGTTCGGCGTCTTTTGATTTAAACGCC  1130
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1462  TACATAACCGGACATAATCATAGGACCTCTCACACACAGTTCGGCGTCTTTTGATTTAAACGCC  1403

Qy   1131  CAGGGTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGGACAACAACTTTACC  1190
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1402  CAGGGTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGGACAACAACTTTACC  1343

Qy   1191  GACCGGCGCCCGGGTTTATCATCCCCCTCGGCGGTGTAATCGAAATAGCTGATGTAGTCTCAGT  1250
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1342  GACCGGCGCCCGGGTTTATCATCCCCCTCGGCGGTGTAATCGAAATAGCTGATGTAGTCTCAGT  1283

Qy   1251  GAGGCCCATATCCTTGCCTGGAACTTCGGCCAACCGGCTTCCCGGAC  1310
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   1282  GAGCCCATATCCTTGCCTGATACCTTGCCGAGATGGAAACCTTCCGGCTTCCCCGGAC  1223

Qy   1311  TTCCTTAGAGAGGGAGAGCCCCACCAGAGCCAATTCGTGTAAATTAGATAAATCGTATTT  1370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222  TTCCTTAGAGAGGGAGAGCGCGCCACCAGAGCAATTCGTGTAAATTAGATAAATCGTATTT  1163

Qy   1371  GTCAATCAGAGTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCCACTTTG  1430
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 
```

```
Db  1162  GTCAATCAGAGTGCTTTTGGCGAAGAAGAGAGAATAGGGTTGGCCACCAGCAGCCACTTTG  1103

Qy  1431  AATCTTGTAATCCTTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGAC  1490
Db  1102  AATCTTGTAATCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGAC  1043

Qy  1491  GACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAA  1550
Db  1042  GACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAA  983

Qy  1551  TGGAACAACACTTAAAATCCCAGTATCCCGAAATGATTTGATTTGCAAAAATAGGATCTCT  1610
Db  982   TGGAACAACACTTAAAATCCCAGTATCCCGAAATGATTTGATTTGCAAAAATAGGATCTCT  923

Qy  1611  GGCATGCGAGAAATCTCACCGCAGGCAGTTCTTATGAGGCAGAGCGACACCTTTAGGCAGACC  1670
Db  922   GGCATGCGAGAAATCTCACCGCAGGCAGTTCTTATGAGGCAGAGCGACACCTTTAGGCAGACC  863

Qy  1671  AGTAGATCCAGAGAGGAGTTCATGATCAGTGCAAATTGTTCTTGTCCCTATCGAAGGACTCTGG  1730
Db  862   AGTAGATCCAGAGAGGAGTTCATGATCAGTGCAAATTGTTCTTGTCCCTATCGAAGGACTCTGG  803

Qy  1731  CACAAAAATCGTATTCATTAAAACCGGGAGGTAGAATGAGAGATGTGACGAACGTGTACATCGA  1790
Db  802   CACAAAAATCGTATTCATTAAAACCGGGAGGTAGAATGAGAGATGTGACGAACGTGTACATCGA  743

Qy  1791  CTGAAATCCCTGGTAAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGGAG  1850
Db  742   CTGAAATCCCTGGTAAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGGAG  683

Qy  1851  CTTTTTTTGCACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGG  1910
Db  682   CTTTTTTTGCACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGG  623

Qy  1911  CTGCCAAATCCCATACTGTTGAGCAATTCACGTTCATTATAAAAGTCGTTCGCGGGGCGC  1970
Db  622   CTGCCAAATCCCATACTGTTGAGCAATTCACGTTCATTATAAAAGTCGTTCGCGGGGCGC  563

Qy  1971  AACTGCAACTCCGATAAATAAACGGGCCCCAACACCACGGGCCATAAAAGAATTGAAGAGAGTTTTC  2030
Db  562   AACTGCAACTCCGATAAATAAACGGGCCCCAACACCACGGGCCATAAAAGAATTGAAGAGAGTTTTC  503

Qy  2031  ACTGCCATACCGACGAATTCTGTGATTTGTATTCAGCCCCATATCGTTTCATAGCTTCTGCCAA  2090
Db  502   ACTGCCATACCGACGAATTCTGTGATTTGTATTCAGCCCCATATCGTTTCATAGCTTCTGCCAA  443

Qy  2091  CCGAACGGACATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATAAGTGCATCTGT  2150
Db  442   CCGAACGGACATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATAAGTGCATCTGT  383

Qy  2151  AAAAGCAATTGTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCGAGTTGCTCTCC  2210
Db  382   AAAAGCAATTGTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCGAGTTGCTCTCC  323

Qy  2211  AGCGGTTCCATCTTCCAGCGGATAGAAATGGCGCCCGGGCCTTTCTTTAATGTTTTTGGCGTC  2270
Db  322   AGCGGTTCCATCTTCCAGCGGATAGAAATGGCGCCCGGGCCTTTCTTTAATGTTTTTGGCGTC  263

Qy  2271  TTCCATGG  2278
Db  262   TTCCATGG  255
```

Search completed: September 29, 2005, 15:00:48
Job time : 1455 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2005, 11:18:47 ; Search time 8765 Seconds
(without alignments)
12033.781 Million cell updates/sec

Title: US-10-734-801-18
Perfect score: 2771
Sequence: 1 ggatccgctggaatgtgt..........atgggaccacacaaatctctc 2771
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match Length	Match DB	DB ID	Description
1	925.4	33.4	1034	8	BZ676934	PUBIF45TD
c 2	914.6	33.0	1032	8	BZ675132	PUB1G87TD
c 3	871.8	31.5	973	8	BZ701200	PUBMP50TD
c 4	861.2	31.1	949	8	BZ681727	PUBH142TD
c 5	852.8	30.8	944	8	BZ672965	PUBGD25TD
c 6	831.6	30.0	902	8	BZ670418	PUBEN17TD
c 7	827.4	29.9	929	8	BZ678861	PUBFK69TD
c 8	811.8	29.3	911	8	BZ671975	PUBEX85TD
c 9	809.8	29.2	885	8	BZ686758	PUBFE30TD
c 10	804.8	29.0	884	8	BZ703458	PUBLO07TD
c 11	803.4	29.0	916	8	BZ679896	PUBEV13TD
c 12	795.6	28.7	866	8	BZ682015	PUBJP02TD
c 13	795.2	28.7	892	8	BZ677942	PUBFU55TD
c 14	793.4	28.6	882	8	BZ682835	PUBGP67TD
c 15	787.6	28.4	867	8	BZ683904	PUBHO53TD
c 16	760.8	27.5	863	8	BZ676173	PUBHT60TD
c 17	748.4	27.0	814	8	BZ685628	PUBFN19TD
c 18	738.8	26.7	822	8	BZ682928	PUBHQ02TD
c 19	735.2	26.5	796	8	BZ669209	PUBCL07TD
c 20	727.8	26.3	811	8	BZ676058	PUBFD85TD
c 21	717.6	25.9	805	8	BZ702755	PUBMR69TD
c 22	716.4	25.9	811	8	BZ687372	PUBHF61TD
c 23	705.8	25.5	789	8	BZ669225	PUBDK64TD
c 24	704.2	25.4	841	8	BZ669574	PUBCO87TD
c 25	696	25.1	796	8	BZ994546	PUBKI46TD
c 26	683.2	24.7	760	8	BZ677712	PUBFP21TD
c 27	632.6	22.8	687	8	BZ683170	PUBC064TD
c 28	603.4	21.8	659	8	CB546382	AMGNNUC:N
c 29	576.2	20.8	637	8	BZ670807	PUBBC62TD
c 30	571.2	20.6	618	8	CB579321	AMGNNUC:N
c 31	536	19.3	574	8	CB608609	AMGNNUC:N
c 32	527	19.0	603	8	BZ700921	PUBL128TD
c 33	515	18.6	855	8	BZ756044	PUBD124TB
c 34	507.4	18.3	541	8	CB612562	AMGNNUC:N
c 35	483	17.4	785	8	BZ648886	PUBFN18TD
c 36	477	17.2	934	8	BZ677304	PUBGB93TD
c 37	464.8	16.8	492	8	BZ684949	PUBCH20TD
c 38	457	16.5	497	8	BZ683468	PUBBD76TD
c 39	449.8	16.2	501	8	BZ701708	PUBMQ56TD
c 40	447.6	16.2	708	8	BY761183	PUBMQ56TD
c 41	370	13.4	527	8	BP538264	BP538264
c 42	367.4	13.3	543	8	BP538266	BP538266
c 43	366.8	13.2	543	8	BP538261	BP538261
c 44	366.2	13.2	546	8	BP538055	BP538055
c 45	363.6	13.1	543	8	BP538265	BP538265

ALIGNMENTS

RESULT 1
BZ676934/c
LOCUS   BZ676934   1034 bp   DNA   linear   GSS 05-FEB-2003
DEFINITION PUBIF45TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa061G18,
            genomic survey sequence.
ACCESSION   BZ676934
VERSION   BZ676934.1 GI:28228896
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 1034)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..1034
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa061G18"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

Query Match   33.4%; Score 925.4; DB 8; Length 1034;
Best Local Similarity   94.0%;
Matches 973; Conservative   0; Mismatches 61; Indels   1; Gaps   1;

Qy   971 CAACGATGAAGAAGTGTTCGTTCTTCGTCCCAGTAAGCTATATGTCTCCAGAAATGAGCCATC 1030
           CAACTATGAAGAAGTGTTCGTC-TCGTCCCAGTAAGCTATATGTCTCCAGAATGTAGCCATC
Db   1034 CAACTATGAAGAAGTGTTCGTC-TCGTCCCAGTAAGCTATATGTCTCCAGAATGTAGCCATC 976

```

```
Qy   1031 CATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCGGACATAATCA 1090
              |||| ||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   975 CATCTTTGTCAATCAAGGCGTTGGTCGCTTCCGGATGTTTTACATAACCGGACATAATCA 916

Qy   1091 TAGGACCTCTCACACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTCCCGGTATCCA 1150
              |||| ||||| ||||| | ||||||||| |||||||||||||||||||||||||||||||
Db   915 TAGGTCCTCTGACACATAATTCGCCTCTCTGATTAACGCCCAGCGTTTTCCCGGTATCCA 856

Qy   1151 GATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCCGGTTTATCAT 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   855 GATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCCGGTTTATCAT 796

Qy   1211 CCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATATCCTTGCCTGA 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db   795 CCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATATCCTTGTCGTA 736

Qy   1271 TACCTGGCAGATGGAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAGAGGGGAGCGC 1330
              | ||||| |||||||| | || |||||||||||||||||||| || || || || ||||
Db   735 TCCCTGGAAGATGGAAGCGTTTTGCAACCGCTTCCCCGACTTCTTTCGAAAGAGGTGCGC 676

Qy   1331 CACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCTTTTGG 1390
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   675 CCCCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCTTTTGG 616

Qy   1391 CGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAATCCTGAAGGC 1450
              ||||||| || ||||||||||| || |||| |||||||||||| |||||||||||||||
Db   615 CGAAGAATGAAAATAGGGTTGGTACTAGCAACGCACTTTGAATTTTGTAATCCTGAAGGG 556

Qy   1451 TCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCA 1510
                | | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   555 ATCGTAAAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCA 496

Qy   1511 AATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAATCG 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   495 AATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAATCG 436

Qy   1571 CAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAGAATCTCACGC 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   435 CAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAGAATCTGACGC 376

Qy   1631 AGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCAGAGGAGTTCA 1690
              ||||||||||||| || || || ||||| ||||| | ||||||||||||||||| ||||
Db   375 AGGCAGTTCTATGCGGAAGGGCCACACCCTTAGGTAACCCAGTAGATCCAGAGGAATTCA 316

Qy   1691 TGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCGTATTCATTAA 1750
              | |||||||||||||| ||||| | ||| ||||||||||| |||||||||||||||||||
Db   315 TTATCAGTGCAATTGTTTTGTCACGATCAAAGGACTCTGGTACAAAATCGTATTCATTAA 256

Qy   1751 AACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCCTGGTAATCCG 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255 AACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCCTGGTAATCCG 196

Qy   1811 TTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGCACGTTCAAAA 1870
              ||||||||||||||||||||||||| ||||| ||||| | |||||||||||||||||||
Db   195 TTTTAGAATCCATGATAATAATTTTCTGGATTATTGGTAATTTTTTTTGCACGTTCAAAA 136

Qy   1871 TTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATGCCCATACTGT 1930
              ||||||||||||||||||||||||| ||||| |||||||||||||||||| ||||||||
Db   135 TTTTTTGCAACCCCTTTTTGGAAACAAACACTACGGTAGGCTGCGAAATGTTCATACTGT 76

Qy   1931 TGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACTCCGATAAATA 1990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75 TGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACTCCGATAAATA 16

Qy   1991 ACGCGCCCAACACCG 2005
              |||||||||||||||
Db   15 ACGCGCCCAACACCG 1

RESULT 2
BZ675132
LOCUS   BZ675132   1032 bp   DNA   linear   GSS 05-FEB-2003
DEFINITION PUBIG87TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa061P05,
```

```
            genomic survey sequence.
ACCESSION   BZ675132
VERSION     BZ675132.1  GI:28225738
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1032)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .1032
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa061P05"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             33.0%;  Score 914.6;  DB 8;  Length 1032;
  Best Local Similarity   93.7%;  Pred. No. 1.6e-176;
  Matches  964;  Conservative    0;  Mismatches   64;  Indels    1;  Gaps    1;

Qy        700 AGACCTTTCGGTACTTCGTCCACAAACACAACTCCTCCGCGCAACTTTTTCGCGGTTGTT 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 AGACCTTTCGGTACTTCGTCCACAAACACAACTCCTCCGCGCAACTTTTTCGCGGTTGTT 64

Qy        760 ACTTGACTGGCGACGTAATCCACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAA 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ACTTGACTGGCGACGTAATCCACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAA 124

Qy        820 ACAACAACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGGAAGACCTGCGACACCTGCG 879
              |||||||||||||||||||||||||||||||||||||||||||||||||| || || |||
Db        125 ACAACAACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGGAAGACCTGCCACGCCCGCG 184

Qy        880 TCGAAGATGTTGGGGTGTTGGAGCAAGATGGATTCCAATTCAGCGGGAGCCACCTGATAG 939
              |||||||||||||||||||| | ||| || ||||||||||||||||| |||||||||||
Db        185 TCGAAGATGTTGGGGTGTTGTAACAATATCGATTCCAATTCAGCGGGGGCCACCTGATAT 244

Qy        940 CCTTTGTACTTAATCAGAGACTTCAGGCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCC 999
              |||||||| ||||| | |||||||| ||||||||| ||||||||||||||||||||||||
Db        245 CCTTTGTATTTAATTAAAGACTTCAAGCGGTCAACTATGAAGAAGTGTTCGTCTTCGTCC 304

Qy       1000 CAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCT 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 CAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCT 364

Qy       1060 TCCGGATTGTTTACATAACCGGACATAATCATAGGACCTCTCACACACAGTTCGCCTCTT 1119
              ||||||||||||||||||||||||||||||||||| ||||| ||||| | |||||||||
Db        365 TCCGGATTGTTTACATAACCGGACATAATCATAGGTCCTCTGACACATAATTCGCCTCTC 424

Qy       1120 TGATTAACGCCCAGCGTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGA 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TGATTAACGCCCAGCGTTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGA 484

Qy       1180 ACAACTTTACCGACCGCGCCCGGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGAT 1239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 3
BZ701200/c

LOCUS   BZ701200   973 bp   DNA   linear   GSS 19-FEB-2003
DEFINITION PUBMP50TD ZM_0_6_1.0_KB Zea mays genomic clone ZMMBTa089J04,
            genomic survey sequence.
ACCESSION   BZ701200
VERSION   BZ701200.1 GI:28421047
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 973)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT   TIGR
            Contact: Cathy Whitelaw
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..973
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa089J04"
                     /clone_lib="ZM_0_6_1.0_KB"
                     /note="Vector:_pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN
```

```
Query Match          31.5%;  Score 871.8;  DB 8;  Length 973;
Best Local Similarity 95.0%;  Pred. No. 9.4e-168;
Matches  900;  Conservative   0;  Mismatches  47;  Indels   0;  Gaps   0;
```

```
Qy   485 ACAACTTTACCGACCGGCGGCCCGGTTTATCATCCCCTCGGGGTGTAATCAGAATAGCTGAT 544
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1240 GTAGTCTCAGTGAGCCCATATCCTTGCCTGATAACCTGGCCAGATGGGAACCTCTTGGCAACC 1299

Qy   545 GTAGTCTCAGTGAGCCCATATCCTTGTCGTATCCCTGGAAGATGGGAACGGTTTTGCAACC 604
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1300 GCTTCCCGACTTCTTCTTAGAGAGGAGCGCGCCACCAGAAGCAATTCGTGTAAAATTAGAT 1359

Qy   605 GCTTCCCGACTTCTTCTTGGAAAGAGGTGCGCCCCCAGAAGCAATTCGTGTAAAATTAGAT 664
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1360 AAATCGTATTTGTCAATCAGAGTGCTTTTGGGCGGAAGAGGAGAATAGGGTTGGCACCAGC 1419

Qy   665 AAATCGTATTTGTCAATCAGAGTGCTTTTGGGCGGAAGAGGAGAATAGGGTTGGCACCAGC 724
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1420 ACGGCACTTTGAATCTTCTAATCCTGAAGGCTCCTCAGAAACAGCTCTCTTCTTCTCAAATCTA 1479

Qy   725 AACGGCACTTTGAATTTTGTAAATCCTGAAGGGTGTAGTTAGTAGTAAAAAACAGCTCTTTGAAT 784
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1480 TACATTAAGACGACTCGGAAATCCACATATCAAAATATCCGAGTGTAGTTAGTAAAAAAATCTA 1539

Qy   785 TACATTAAGACGACTCGGAAATCCACATATCAAAATATCCGAGTGTAGTTAGTAAAAAATTCAAA 844
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1540 CCGTGATGGAATGGAACACACTTAAAATCGCAGTATCCGGAATGCAGTAGTTGATTTGAATGCAGCA 1599

Qy   845 CCGTGATGGAATGGAACAACACTTAAAATCGCAGTATCCGGAATGATTTGAATTGCCAAAA 904
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1600 ATAGGACATCTCGCATGCGAGAATCTCACGCAGCAGTCTTATGAGCGAGAGGCACCT 1659

Qy   905 ATAGGATCTCTGGCATGCAGGAATCTGACGCAGTTCTATGCGGAAGGGCCCACACC- 963
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1660 TTAGGCAGACCAGTAGATCCAGGAGGTTCATGATCAGTGCAAATTGTTGTTCTTATTCG 1719

Qy   964 CTAGGTAACCCAGTAGATCCAGGAGGAATTCATATCAGTGCAAATTGTTTTTGTCACGGATCA 1023
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1024 AAGGACTCT 1032
```

```
Qy  1024 AAGGACTCT 1032
         |||||||||
Db  1720 AAGGACTCT 1728
```

```
Qy  1330 CCACCGAGAGCAATTTCCTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCTTTTTG 1389
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCCCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCTTTTTG 902

Qy  1390 GCGGAAGAGGAGAATAGGGTTGGCACCAGCACGGCGCACTTTGAATCTTGTAAATCCTGAAGG 1449
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GCGGAAGAATGAAAATAGGGGTTGGTACTAGCAACGCACTTTGAATTTGTAAATCCTGAAGG 842

Qy  1450 CTCCTCAGAAACAGCTCTTCTTCTTCAAAATCTATACATTAACATTACAATCCACATATC 1509
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GATCGTAAAAACAGCTCTTTCTTCTTCAAAATCTATATACATTAAGACGACTCGGAAATC 782

Qy  1510 AAATATCCGAGTGTAGTTAGTAAAAATTCCAAAACATTCCAAAATCTATACATTAAGATC 1569
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AAATATCCCAGTGTGTAGTTAGTAAAACATTCCAAAACCGGTAACGGTAGAATCGAAAATC 722

Qy  1570 GCAGTATCCGGAATGATTTGAATTGCCAAAAATAGGACTCTCTGGCATGCGAGGAATCTCACG 1629
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCAGTATCCGGGAATGATTTGAATTGCCAAAAATAGGAATCTCTCGCGGATGCGAGAATCTGACG 662

Qy  1630 CAGGCAGTTCTATGAGCGAGCAGCACACTTTAGGCAGACCAGTAGATCCAGGAGGTTC 1689
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CAGGCAGTTCTATGCGGAAGGGCCACACCCTTAGGTAACCCAGTAGATCCAGGAGGAATTC 602

Qy  1690 ATGATCAGTGCAAATTGTTCTTGTCCCTATCGAACGGAAGGGACTCTGGCACAAAAATCGTATTCATTA 1749
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ATTATCAGTGCAAATTGTTTTTGTCACGGACTCTGGTACAAAAATCGTATTCATTA 542

Qy  1750 AAACCGGAGGTAGAATGAGATGAGAGTGAACGGAACGGTGTACAATCGACTGAAAATCCCTGGTAATCC 1809
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAACCGGAGGTAGAATGAGAATGAGATGAACGGTACAATCGACTGAAAATCCCTGGTAATCC 482

Qy  1810 GTTTTAGAATCCATGATAATAAATTTTTTTTTTGGATGATTGGGAGCTTTTTTTTGCACGTTCAAA 1869
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GTTTTAGAATCCATGATAATAAATTTTCTGGATTTATGGTAATTTTTTTTGCACGTTCAAA 422

Qy  1870 ATTTTTTTTGCAACCCCTTTTTGGAAAACGAACACACAGGTAGGCTGCGAAATGCCCATACTG 1929
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ATTTTTTGCAACCCCTTTTTGGAAACACTACGGTAGGCTGCGAAAGTTCATACTG 362

Qy  1930 TTGAGCAATTCAGTTCATTATAAAATGCGTTGCGGGGGGCCAACTGCAACTCCGGATAAAT 1989
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TTGAGCAATTCACGTTCATTATAAAATGTCGTTGCGGGGGCCAACTCGGCAACTCCGGATAAAT 302

Qy  1990 AACGCGCCCAACACCGGCGTAAAGAGAATTGAAGAGAGTTTTCACTGCATACGACCATTCTG 2049
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AACGCGCCCAACACCGGCGATAAAGAATGGAAGAGAGTTTTCACTGCATACGACGATTCC 242

Qy  2050 TGATTGTGTAATTCAGCCCATATCGTTTCATTAGCCTTCTCTGCCAACCGAACCGGACATTCGAAG 2109
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGATTGTGTAATTCAGCCCATATCGTTTCATTAGCTTCCTGCCAACCGGACATTCGAAG 182

Qy  2110 TACTCAGGGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGA 2169
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TATTCCGCGGTACGTGATGTTGCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGA 122

Qy  2170 ACCAGGGGCGTATCTCTTCATTAGCGTTGCTCTCCACGGGTTCCATCTTCCCAGC 2229
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACCAGGGGGGTATCTCTTCATTAGCCTTAGTGTTGCTCTTCCCAGCCGGTTCCATCTCCTGA 62

Qy  2230 GGATAGAATGGGGGCCGCCGGGCCCTTCTTTCTTATAGTGTTTTTTTGGGGTCTTCCAT 2276
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GGATAGAATGGCGACGGGCCCTTTCTTTAGTGTTTTTATGGGGTCTTCCAT 15
```

RESULT 4
LOCUS   BZ681727
DEFINITION   BZ681727/c
                PUBHI42TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa055H11,
                genomic survey sequence.
ACCESSION   BZ681727   949 bp   DNA   linear   GSS 05-FEB-2003
VERSION   BZ681727.1 GI:28237508
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM   Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 949)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
                Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
                Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA
                Tel: 301-838-5843
                Fax: 301-838-0208
                Email: whitelaw@tigr.org
                Seq primer: TF
                Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..949
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa055H11"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

Query Match   31.1%; Score 861.2; DB 8; Length 949;
Best Local Similarity 94.9%; Pred. No. 1.4e-165;
Matches 890; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy   1344 TTCGGTGTAAATTAGATAAATCGTATTGTTGTCAATCAGAGTGCTTTTGGGGAAGAGGAGAA 1403
            |||||||||||||||| |||||||||||||||||| ||||||||||||||||| |||||||
Db   949 TTCGTGTAAATTAGATAAATCGTATTTGTTGTCAATCAGAGTGCTTTTGGGGCCGAAGATGAAAA 890

Qy   1404 TAGGGTGGCCACCAGCAGCGCACTTTGAATCTTGTATCCTGAAGGCGTTCCTCAGAAACAG 1463
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   889 TAGGGTTGGTACTAGCAAGCGCACTTTGAATTTGTTAATCCTGAAGGGATCGTAAAAACAG 830

Qy   1464 CTCCTCTTCAAAATCTATTACATTAAGACGATCTCGAAAATCCACATATCAAATATATCCGAGTGT 1523
            ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   829 CTCTTCTTCAAAATCTATTACATTAAGACGATCTCGAAAATCCACATATCAAATATATCCGAGTGT 770

Qy   1524 AGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAATGGCAGTATCCGAGAAT 1583
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   769 AGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAATGGCAGTATCCGAGAAT 710

Qy   1584 GATTTGATTGCCAAAAATAGGAGACTCTCTGCCATCGAGGAATCTCACGCACGCACTTTCTATG 1643
            ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   709 GATTTGATTGCCAAAAATAGGAATCTCTGCCATGGCAGGCAGGCAGTCTATG 650

Qy   1644 AGGCAGAGGAGGGACGACACCTTTAGGCAGAGCAGTAGATCCAGGGAGTTCATCGATCGATGCAAT 1703
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   649 CGGAAGGGCCCACACCCTTAGGTAACCCAGTAGATCCAGAGAATTCATTATCAGTGCAAT 590

Qy   1704 TGTGTTGTCCCTATCGAAGGACTCTGGGACACAAAATGTATTCATTAAAAACCGGGACGGGTTAG 1763
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   589 TGTTTTGTCACGGATCAAAAGGACTCTGGTGTACAAAAATCGTATTCATTAAAAACCGGGAAGGTAG 530

Qy   1764 ATGAGAGTGACGAACGTGTACATCGACTGAAAATCCCGTTTAGAATCCAT 1823
            |||||||||||||||||||||||||||||||||||||||||||||
Db   12

RESULT 5
LOCUS   BZ672965
DEFINITION   BZ672965/c
                PUBGD25TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa047F02,
                genomic survey sequence.
ACCESSION   BZ672965   944 bp   DNA   linear   GSS 05-FEB-2003
VERSION   BZ672965.1 GI:28221885
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM   Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 944)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
                Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
                Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA
                Tel: 301-838-5843
                Fax: 301-838-0208
                Email: whitelaw@tigr.org
                Seq primer: TF
                Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..944
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa047F02"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"

Db   529 ATGAGAGTGACGAACGTGTACATCGACTGAAAATCCCGTTTAGAATCCAT 470

Qy   1824 GATAATAAATTTTTTGGATGATTGGGAGGCTTTTTTTGTGCAGTTCAAAAATTTTTTGCAACCC 1883
            |||| ||||||||| ||||||||||||||||| ||||||||| ||||||||||||| |||
Db   469 GATAATAAATTTCTGGATTATTGGTAATTTTTTGCACGTTCAAAATTTTTTGCAACCC 410

Qy   1884 CTTTTTTGGAAACGAACACACCACGTAGGCTGCGGAAATGCCCATACTACTTGTTGAGCAATTCACG 1943
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   409 CTTTTTTGGAAACAAACACTACGGTGAGGCTGCGGAAATGTTCATACTGTTTGAGCAATTCACG 350

Qy   1944 TTCATTATTAAAATGTGCTTCCGCGGGGCGCAACTGCGACTCGGATAAAATAAACCGCGCCCAACAC 2003
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   349 TTCATTATTAAAATGTCGTTCGCGGGGCGCAACTCCGGATAAAATAACCGCGCCCAACAC 290

Qy   2004 CGGCATAAAGAAAATTGAAGACAGAGTTTTCACTGCATACGACGGATTCGTTGATTGTTGATTCAG 2063
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289 CGGCATAAAGAAATTGAAGAGAGTTTTCACTGCATACGACGATTCTGTGATTTGTTGATTTCAG 230

Qy   2064 CCCATATCGTTTCATTAGCTTCTGCCAACCGACAACCGACACATTTCGAAGTAGTACGTTCAGCGTAAGT 2123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229 CCCATATCGTTTCATTAGCTTCTGCCAACCGACAACCGACACATTTCGAAGTAGTACTTCAGCGTAAGT 170

Qy   2124 GAGTCCACCTCGATATGTGCATCGTAGTAAAAGCAAATTGTTCCAGGAACACAGGGCGTATCT 2183
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   169 GATGTTCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGAACACCAGGGCGTATCT 110

Qy   2184 CTTCATAGCCTTAGTGCAGTTGCTCTTCCCAACGGGTTCCATCTTCCAGCGCGGTAGCGGAGAATGGGGC 2243
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   109 CTTCATAGCCTTATCGCAGTTGCTCTTCCCAGCGGGTTCCATCCCTCTAGGAGGATAGAAATGGGGC 50

Qy   2244 CGGGGCCTTTCTTTTAAGTTTTTTTGGGGTCTTCCCATTTTAC 2281
            ||||||||||||||||||||||||||||||||||||||||||
Db   49 CGGGGCCTTTCTTTAAGTTTTTTTGGGGTCTTCCCATTTTAC 12

RESULT 6
BZ670418/c

Left alignment (Query vs BZ670418)

ORIGIN
Best Local Similarity   30.8%;   Score 852.8; DB 8;   Length 944;
Query Match   94.0%;   Pred. No. 7.2e-164;
Matches 887; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy   1062 CGGATTGTTTACATAACCGGACTATAATCATTAGGACTTCTCACACACAGTTCGCCTCTTTG 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 CGGATTGTTTACATAACCGGACTATAATCATTAGGTCCCTCTGACACATAATTCGCCTCTG  885

Qy   1122 ATTAACGGCCCAGGCGTTTTCCCGGTATCCAGATCACCAACCTTGGCTTCAAAAAATGGAAC 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 ATTAACGCCCAGGCGTTTTCCCGGTATCCAGATCCAACCTTCGCTTCAAAAAATGGAAC  825

Qy   1182 AACTTTACCGACCGCCGCCGGTTTATCATCCCCTCGCGGTGTAATCAGAATAGCTGATGT 1241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    824 AACTTTACCGACCGCCGCCGGTTTATCATCCCCTCGGCGTGTAATCAGAATAGCTGATGT  765

Qy   1242 AGTCTCAGTGAGCCCATATCCTTGCCTCAGATGAGAAACTTCTTGGCAACCGC 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 AGTCTCAGTGAGCCCATATCCTTGTCGTATCCCTGGACGAAGCGTTTTGCAACCGC  705

Qy   1302 TTCCCCGACTTCCTTAGAGAGGGGGGAGCGCGCCACCAGAGAGCAAATTCGTGTAAATTAGATAA 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    704 TTCCCCGACTTCTTTCGAAGAGAGGTGCGCCCCCAGAAGCAAATTCGTGTAAATTAGATAA  645

Qy   1362 ATCGTATTTGTCAATCAGAGTGCTTTGGCGAAGAAGGAGAATAGGGGTTGGGCACCAGCAG 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    644 ATCGTATTTGTCAATCAGAGTGCTTTGGCGAAGAAATGAAAAATAGGGGTTGGTACTAGCAA  585

Qy   1422 CGCACTTTGAAATCTTGTAATCTTCTTCGAAGGCTTCCTCAGAAACAGCTCTTCTTCTCAAATCTATA 1481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    584 CGCACTTTGAATTTGTAATCCTGAAGGGATCGTAAAAAACAGCTCTTCTTCTCAAATCTATA  525

Qy   1482 CATTAAGAGACGACTCGAAAATCCACATATCAAAATATCCGAGTGTAGTGAACATTCCAAAACC 1541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    524 CATTAAGAGACACTCGAAAATCCACATATCAAAATATCCGAGTGTAGTGAACATTCCAAAACC  465

Qy   1542 GTGATGGAATGGGAACAACACTTAAAATCGCAGTATCCGAATGATTTTGAATTGCCAAAAAT 1601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    464 GTGATGGAATGGGAACAACACTTAAAATCGCAGTATCCGGAATGATTTTGATTGCCAAAAAT  405

Qy   1602 AGGGATCTCTTGGCATGCGAGAATCTCACGCAGGCCAGTTCTATGAGGCAGAGCGACCACTTT 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 AGGATCTCTTGGCATGCGAGAATCTGACGGCAGTTCTTATGTCGGAAGGGCCACACCCTT  345

Qy   1662 AGGCGAGACCAGTAGAATCCAGAGAGGAGTTCATGATCAGGTCGAAATTGTCTTGTCCCTATCGAA 1721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 AGGTAACCCAGTAGAATCCAGAGGAATTCATTATCAGTGCAATTGTTTTGTGTCACGGATCAAA  285

Qy   1722 GGACTCTCTTGGCACAAAAATCGTATTCATTAAAAAACCGGGAGGTTAGAGATGTGAGGAACGT 1781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    284 GGACTCTGTGTACAAAAATCGTATTCATTAAAAACCGGGTAGAGTGAGAGTGAGGAAACGT  225

Qy   1782 GTACATCGACTGAAATCCCTGGTAATCCGTTTTAGAATCCATGAATAATAATTTTTTGGAT 1841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    224 GTACATCGACTGAAATCCCTGGTAATCCGTTTTAGAATCCATGAATAATAATTTTCTGGAT  165

Qy   1842 GATTGGGGACTTTTTTTTTTTTGCACGTTCAAAAATTTTTTGGAAAACGAACAC 1901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    164 TATTCGGTAATTTTTTTTGCACGTTCAAAAAATTTTTTGGAAAACAAAACAC  105

Qy   1902 CACCGGTAGGCTGCGAAAATGCCCATACTGTTGAGCAATTCACGTTCAGTTCATTATAAATGTCGTT 1961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104 TACGGTAGGCTGCGAAAATGTTCATACTGTTGAGCAATTCACGTTCATTATAAAGTGTCGTT 1961

Qy   1962 CGGCGGGGGCGAACTGCAACTCCGATAAATAACGGCGGCCAACACACCG 2005
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db     44 CGGCGGGGGCGAACTCGATAAATAACGGCGGCCAACACCG 1
```

Right alignment (Query vs BZ670418 — RESULT 6)

Best Local Similarity   30.0%;   Score 831.6; DB 8;   Length 902;
Query Match   95.1%;   Pred. No. 1.6e-159;
Matches 858; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy   1344 TTCGTGTAAATTAGATAAAATCGTATTTGTCAATCAGAGTGCTTTGGCGAAGAAGGAGAA 1403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    902 TTCGTGTAAATTAGATAAAATCGTATTTGTCAATCAGAGTGCTTTGGCGAAGAAGGAGAA  843

Qy   1404 TAGGGTTGGCACCAGCAGCGCACTTTGAAATCTTGTAATCTTGTAATCTCAGAGAACAG 1463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    842 TAGGGTTGGGTACTAGCAACCGCACTTTGAATTTGTAATCCTGAAGGATCGTAAAAAACAG  783

Qy   1464 CTCTTCTTCAAATCTATACATTAAGAGACGACTCGAAAATCCACATATCAAAATATCCGAGTGT 1523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    782 CTCTTCTTCAAATCTATACATTAAGAGACACTCGAAAATCCACATATCAAAATATCCGAGTGT  723

Qy   1524 AGTAAACATTCCAAAACCGTGATGGAATGGGAACAACACACTTAAAATCGCAGTATCCGGAAT 1583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    722 AGTAAACATTCCAAAACCGTGATGGAATGGGAACAACACTTAAAATCGCAGTATCCGGAAT  663

Qy   1584 GATTTTGATTGCCAAAAATAGGGATCTCTTGGCATGCGAGAATCTCAGGCAGCGCAGTTCTATG 1643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 GATTTTGATTGCCAAAAATAGGATCTCTTGGCATGCGAGAATCTGACGGCAGTTCTATG  603

Qy   1644 AGGCAGAGCGACCACTTTAGGCGAGACCAGTAGAATCCAGAGGAGTTCATGATCAGTGCAAT 1703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 CGGAAGGGCCACACCCTTAGGTAACCCAGTAGAATCCAGAGGAGTTCATTATCAGTGCAAT  543

Qy   1704 TGTCTTGTCCCTATCGAAGGACTCTCTTGGCACAAAAATCGTATTCATTAAAAAACCGGGTAG 1763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 TGTTTTGTGTCACGGATCAAAGGACTCTGTGTACAAAAATCGTATTCATTAAAAACCGGGTAG  483

Qy   1764 ATGAGATGTGACGAACGTGTACATGACTGAAATCCCTGGTAATCCGTTTTTAGAATCCAT 1823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 ATGAGATGTGACGAACGTGTACATGACTGAAATCCCTGGTAATCCGTTTTTAGAATCCAT  423
```

```
Qy    1824  GATAATAATTTTTTGGATGATTGGGAGCTTTTTTTTTGGCAGTTCAGTTCAAAAATTTTTTTGCAACCC  1883
            ||||||||||||  |||||| ||||| |||||||||| |||||||   |||||||  |||||||||||||
Db     422  GATAATAATTTTCTGGATTATTGGTAATTTTTTTTGGCAGTTCAAAAATTTTTGCAACCC           363

Qy    1884  CTTTTTTGGAAACGAACACACCACGGTAGCTGCGGAAAATGCCCATACTGTTGAGCAATTCACG       1943
            |||||||||||| ||||| || || ||||| |||| |||||| ||||||||||||||||||||
Db     362  CTTTTTTGGAAACAAACACTACCGGTAGGCTGCGAAAATGTTCAATACTGTTGAGCAATTCACG       303

Qy    1944  TTCATTATATAAAGTGTCGTTCGCGGGGCGGGGCGGCAACTGCAACTGCCCCAACAC             2003
            |||||||| || |||||||||||||||| |||| |||||||||| |||||||||||||
Db     302  TTCATTATAAATGTCGTTCGCGGGGCCGGGCCGGCAACTGCACCTGCCCCAACAC               243

Qy    2004  CGGCATAAAAGAATTGAAGAGAGTTTTCACTGCATTCTGTGATTTTTGATTTGTTATTCAG         2063
            |||||||||| ||| ||||||||||||||| |||||| |||||||||| ||||||||||||
Db     242  CGGGCATAAAAGAATTGAAGAGAGTTTTCACTGCATTCTTGTGATTTTTTGATTTGTTATTCAG      183

Qy    2064  CCCATATCGTTTTCATAGCTTCTCTGCCAACCGAAGACATTTCGAAGTACTCAGGCGTAAGT        2123
            ||| |||||||||||||||||||| || |||||||||||||| ||||||||||| ||||| |
Db     182  CCCATATCGTTTCATAGCTTCTCTGCCAACCGAACGGACATTTCGAAGTATTCCGCGTACGT        123

Qy    2124  GATGTCCACCTCGATATGTGATACTGTAAAAGCAATTGTTCCAGGACCACAGGGCGTATCT         2183
            ||||||||||| ||| |||| |||| ||||||| |||||||||||| |||| |||| ||||
Db     122  GATGTTCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGACCCAGGGCCGGTATCT        63

Qy    2184  CTTCATTAGCCCTTATGCCGGTTCATCTTCCAGGGGGATAGAAATGGCGGC                   2243
            ||||||| |||| ||| ||||||||| ||| |||||||||| ||||| ||||
Db      62  CTTCATAGCCTTATGCAGTTGCTCTCCAGAGGATAGAAATGGGCGGC                       3

Qy    2244  CG  2245
            ||
Db       2  CG  1
```

Query Match 29.9%; Score 827.4; DB 8; Length 929;

Best Local Similarity 93.9%; Pred. No. 1 1e-159;
Matches 872; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

```
Qy     747  TTTCGGCGGTTGTTACTTGACTGGCGACGTAATCCACGATCTCTTTTTCCGTCATCGTCTT  806
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  TTTCGGCGGTTGTTACTTGACTGGCGACGTAATCCACGATCTCTTTTTCCGTCATCGTCTT  60

Qy     807  TCCGTGCTCCAAAACAACAACACGCGGC-GGGAAGTTCACCGGCGTCGGGAGAC         865
            ||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db      61  TCCGTGCTCCAAAACAACAACACGCGGCGGGGAAGTTCACCGGCGTCGGGAGAC          120

Qy     866  CTGCGACACCTCGGTCGAAGATGTTGGGGTGTTGGGGGAAGTTCACGGGG      925
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  CTGCCACCGCCCCGGGTCGAAGATGTTGGGGTGTTGTTGAACAATTCACGGG           180

Qy     926  GAGCCACCTGATAGCCTTTGTACTTAATCAGAGGACTTCAGGGGTCAACGGATGAAGAGT   985
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  GGGCCACCTGATATCCTTTGTATTTAATTCAGAGACTTCAAGCGGTCAACTATGAAGAAGT  240

Qy     986  GTTCGTCTTCGTCGAGTAAGCTATGTCTCCAGAAGTTCAGGAATCTGTCAATCA       1045
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  GTTCGTCTTCGTCTCCCAGTGAGCTATGTCTCCAGAAAGTTGCAGATCCATCCTTTGTCACAC 300

Qy    1046  AGGCGGTTGGTGTTCCCGGATTGTTTCATAACCGGACACATAATCATAGGAC        1105
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  AGGCGGTTGGTCGCCTTCCGGATTGTTTACATAACCGGACACATAATCATAGGACCTTCTCTCACAC 360

Qy    1106  ACAGTTCGGCCTCTTTGATTAACGCCCCAGCCGTTTTCCCGGTATCCACGAATCC       1165
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  ATAATTCGGCCCCTCTCTGATTAACGCCCACCGGTATTCCCGGGATATCCACAACCTTCG    420

Qy    1166  CTTCAAAAAATGGAACAAACTTTACCGACACGGCCCCCCCGTTTATCATCCCCCTCGGGGTC 1225
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421  CTTCAAAAAATGGAACAAACTTTACCGACACGGCCGGTTTATCATCCCCCTCGGGGTGTAA  480

Qy    1226  TCAGAATAGCTGATGTAGTGTTCTCAGTGAGCCCCATATCCTTGCCTGATACCGTGGCAGATGGA 1285
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  TCAGAATAGCTGATGTAGTGTTCTCAGTGAGCCCCATATCCTTGCCTGATACCGTGGCAGATGGA 540

Qy    1286  ACCTCTTGGCAACCGCGTTCCCGACTTCCTTAGAGAGGGGGGAGCCCACCAGAGACCAATT 1345
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541  AGCGTTTTGCAACCGCTTCCCGACTTCCTTTCGGAAGGTGCGCCCCCAGAAGCAATT     600

Qy    1346  CGTGTAAATTAGATAAAATCGTATTTCTGTTCAATCAGAGTGCTTTTGGCGAGGAAGGAGAATA 1405
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601  CGTGTAAATTAGATAAAATCGTATTTCTGTTCAATCAGAGTGCTTTTGGGCGAAGGAAGTGAAAATA 660

Qy    1406  GGGTTGGCACCAGCAGCCGCACATTTGAATCTTGTAATCCTCAGAAGGCTCCTCAGAAACAGCT 1465
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661  GGGTTGGTACTAGCAACGCACTTTGAATTTTGTAATCCTGAAAGGGGGATCGTAAAAACAGCT   720

Qy    1466  CTTCTTCAAAATCTATACATTAAGACGGACTCGAAATCCACATATCAAATATCGGAGTGTAG 1525
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721  CTTCTTCAAAATCTATACATTAAGACGGACTCGAAATCCACATATCAAATATCGGAGTGTAG    780

Qy    1526  TAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAAATCGCGAATAGGAATGA     1585
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781  TAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAAATCGCGAGTTCTATGGCG    840

Qy    1586  TTTGATTGCCAAAATAGAGATCTCTGGCATGCGACGAATCTCACCGCAGGCGAGTTCTATGAG 1645
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  TTTGATTGCCAAAAATAGGATCTCTGACGAATCTGACGACCGCAGGCGAGTTCTATTCG    900

Qy    1646  GCAGAGCGACGACCACCCTTAGGCAGGACCAGTA  1674
            |||||||||||||||||||||||||||||||||
Db     901  GAAGGGGCCACACCCCTAGGTAACCCAGTA      929
```

ACCESSION   BZ671975
VERSION   BZ671975.1   GI:28220321
KEYWORDS   GSS.
SOURCE   Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 911)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
TITLE   Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   TIGR
            Contact: Cathy Whitelaw
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..911
                     /organism="zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa039P02"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

Query Match   29.3%; Score 811.8; DB 8; Length 911;
Best Local Similarity 93.2%; Pred. No. 1.8e-155;
Matches 849; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy   845 CGGGGTCATCGTTCGGGAAGAACCCTCGGACACACCTGCTCGTCGAAGAGATGTTGGGCA 904
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGGGGTCATCGTTCGGGAAGAACCTGCCACGCCCGTCGGAAGAGATGTTGGGCGTGTGTAACA 60

Qy   905 AGATGGATTCAATTCAGTCACGCGGGAGCCGCCACCTGATAGCCTTTGTTACTTAATCAGAGACTTCA 964
              | | | ||||||||||||||||||||||||||||||||||||||||||| |||||| ||||||
Db   61 ATATCGGATTCCAATTCAGCGCGGGCCACCTGATATCCTTTGTATTATTAAATTAAAGACTTCA 120

Qy   965 GGCGGTCACGGATGAAGAGAACTGTTTCGTCTTTCGTCTCCCAGRAATGTA 1024
              | ||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGCGGTCAACTATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCCTATGTCTCCAGAATGTA 180

Qy   1025 GCCATCCATCCTTGTCAAATCAAGGCGGTTGGTCGCTTCCGGAATTGTTTACATAACCGGACA 1084
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GCCATCCATCCTTGTCAATCAAAGGGGTTGGTCGCTTCCGGATTGTTTACATAACCGGACA 240

Qy   1085 TAATCATAGGACCTCTCACACACAGTTCGCCCTCTTTGAATTAACGCCCAGCGTTTCCCG 1144
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TAATCATAGGTCCTCTGACACATAATTCGGCTCTCTGAATTAACGCCCAGCGGTTTCCCG 300

Qy   1145 TATCCAGATCCACACACTTCGGCTTCGAGAACAACTTACCGACCGGCCCGGTT 1204
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TATCCAGATCCACAACCTTCGCTTCAAAAAAATGGACAACTTACCGACCGGCCGGTT 360

Qy   1205 TATCATCCCCTCCGGGTGTAATCAGAATAGCTGAATGGTAGTAGTCTCAGTGAGCCCATATCCTT 1264
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TATCATCCCCTCCGGGTGTAATCAGAATACCTGAGTGTAGTGAGTGTTCACTGAGCCCATATCCTT 420

Qy   1265 GCCTGATACCTTGGCAGATGGGAAACCTCTTGGCCAACCCGCTTCCCTTAGAGAGGG 1324
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTCGTATCCCTGGAAGGATGGAAGCGTATTTGCAACCGGTGTCCCGGACTTCTTCTTCGAAAGAG 480

Qy   1325 GAGCGCCCCACCAGGAGCACAATTTCGTGTTAAAATTAGATAAAATCGTATTTGTTCAATCAGAGTGC 1384
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GTGCGCGCCCCAGAAGCGAAATTTCGTGTTAAAATCCGTATTTGTCAATCAGAGTGC 540

RESULT 9
LOCUS   BZ686758   885 bp   DNA   linear   GSS 05-FEB-2003
DEFINITION PUBFE301D_ZM_0.6_1.0_KB_zea mays genomic clone ZMMBTa041F11,
            genomic survey sequence.
ACCESSION   BZ686758
VERSION   BZ686758.1   GI:28247291
KEYWORDS   GSS.
SOURCE   Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 885)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
TITLE   Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   TIGR
            Contact: Cathy Whitelaw
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..885
                     /organism="zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa041F11"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

Query Match   29.2%; Score 809.8; DB 8; Length 885;
Best Local Similarity 94.7%; Pred. No. 4.5e-155;
Matches 838; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy   1385 TTTTGGCGAAGAAGAGAGAATAGGGGTTGGGCACCAGCAGCGCAGCGCGCGCACTTTGAATCTTGTAATCCT 1444
              |||||||||||||| || | |||||||||||||||||||||||||||||||||||||||||||
Db   541 TTTTGGCGAAGAAATGAAAAATAGGGGTTGGGTACTAGCAGCAAGCGCACTTTGAATTTTGTAATCCT 600

Qy   1445 GAAGGGCTCCTCAGAAAACAGCTCTTCCTTCTTCAAATCTATACATTATACAGACGACTGAAATCCAC 1504
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GAAGGGAATCGTAAAACAGCGCTTCTTCAAATCTATACATTAAGACGACTGAAATCCAC 660

Qy   1505 ATATCAAAATATATCCGAGTGTTAGTGAATCAAATCCAAAAACCGTGATATGGAGATGGAACAACACTTA 1564
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATATCAAAATATCCGAGTGTAGTGAAACATTCCAAAAACCGTGATGGAATGGAACAACACTTA 720

Qy   1565 AAAATCCCAGTATCCGGAAATGATTTGATTGCCAAAAAATAGGATCTCTGGAGCGAGGAATC 1624
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   721 AAAATGGCAGTATCCGGAATGATTTGATTGCCAAAAAATAGGATTCTCTGGCATGCGAGGAATC 780

Qy   1625 TCACGCAGGCAGTTCTATGAGGCAGGAGGCGACACCTTTAGGCAGACGTAGATCCAGAGG 1684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TGACGCAGGGCAGTTCTTATGCGGGAAGGGCCACACCCTTAGGTTAACCCAGTAGATCCAGAGG 840

Qy   1685 AGTTCATTGATCAGTGCGCAAATTGTCTTCTTGTCCCTTATCCGAAGGACTCTCGCACAAAATCGTATT 1744
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AATTCATTATCAGTGCAAATTGTTTTGTCACGATCAAAAGGACTCTGGTACAAAAATCGTATT 900

Qy   1745 CATTAAAAACCG 1755
              ||||||||||||
Db   901 CATTAAAAACCG 911

```
RESULT 10
LOCUS       BZ703458             884 bp    DNA     linear   GSS 19-FEB-2003
DEFINITION  BZ703458 PUBLOO7TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa083A13,
            genomic survey sequence.
ACCESSION   BZ703458
VERSION     BZ703458.1  GI:28423866
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 884)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..884
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa083A13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN
Query Match       29.0%;  Score 804.8;  DB 8;  Length 884;
Best Local Similarity   94.7%;  Pred. No. 4.8e-154;
Matches  833;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;
```

```
Qy        695 CGGTAAGACCTTTCGGTACTTCGTCGTCCACAAACACAACTCAACTCCTCCGGCGGCGAACTTTTTCGGCGG 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGTAAGACCTTTCGGTACTTCGTCGTCCACAAACACAACTCAACTCCTCCGGCGGCGAATTTTTTCGGCGG 60

Qy        755 TTGTTACTTGACTGGGCGACGTAATCCACGATCTCTTTTCCGTCATCGTCTTTCCGTGCT 814
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCGTCGAAGAGTGGGGTGTTGTTGACAATGAAGAAGTTCAGCGTCGTTCGGGGGGCCACCT 120

Qy        815 CCAAAACAACAACGACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGCGGAAGACCTGCGACAC 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAAAAACAACAACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGCGGAAGACCTGCCACGC 180

Qy        875 CTGCGCTCGAAGATGTTGGGGTGTTGGGAGCAAGAGATGGATTCCAAATTCAGCGGAGCCACCT 934
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCGGTCGAAGAGTGGGGTGTTGTTGACAATGAAGAAGTTCAGCGTCGTTCCGGGGGCCACCT 240

Qy        935 GATAGGCTTTCGTACTTAATCAGAGACTTCAGCGGCGTCAACGATGAAGAAGTGTTCGTGCTT 994
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATATCCTTTGTATTTAATTAAAGACTTCACCGGCGTCAACTATGAAGAAGTGTTCGTCTT 300

Qy        995 CGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTCCTTGTCAATCAAGGCGTTGG 1054
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTCCTTGTCAATCAAGGCGTTGG 360

Qy       1055 TCGGCTTCCCGGATTGTTGTTTTACATAAACCGGACATAATCATTAGGAGCACCTTCTCACACACAGTTCGC 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGCTTCCCGGATTGTTGTTTTACATAAACCGGACATAATCATTAGGGTCCTCTTCGACACACATAATTCGC 420

Qy       1115 CTCTTTGATTAAACGCCCAGCGGTTTTCCCGGTATCCACAGATCCAACACCTTCGCTTCAAAAA 1174
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCTCTGATTAACGCCCAGCGTTTTCCCGGTATCCAACACCTTCGCTTCAAAAA 480

Qy       1175 ATGGAACAACTTTACCGACACCGCCGCCGGGTTTATCATCCCCCTGCGGTGTTAATCAGAATAG 1234
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGAACAACTTTACCGACACCGCCGCCGGGTTTATCATCCCCCTGGGTGTTAATCAGAATAG 540

Qy       1235 CTGAGTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTGGCATGGAACCTCTTGG 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGATGTAGTCTCAGTGAGCCCATATCCTTGTCTGGCATGGAAGGCGTTTG 600

Qy       1295 CAACCGCTTCCCCGACTTCCTTAGAGAGGGGAGCGCCACCAGAGAGCCAATTTCGTGTGTAAAT 1354
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAACCGCTTCCCGACTTCTTTCGAAAGAGGTGCGCCCCAGAAGCCAATTTCGTGTGTAAAT 660

Qy       1355 TAGATAAATGTGTATTTGTCAAATCAGAGTGCTTTTGGCGAAGAAGAGAGATAGGGTTGGGCA 1414
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAGATAAATGTGTATTTGTCAAATCAGAGTGCTTTTGGCGAAGAATGAAAATAGGGTTGGTA 720

Qy       1415 CCAGCAGCGGCACTTTGAAATCTTGTGTAATCCTGAAGGCTCCTCAGAAAAACAGCTCTTCTTCAA 1474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTAGCAAGCGGCACTTTGAAATTTGTATTCCTGAAGGGATCGTTAAAAACAGCTCTCTTCTTCAA 780

Qy       1475 ATCTATACATTAAGAGACGACTCGGAAATCCACATATCAAATATCCGAGTGTAGTGTAAACATTC 1534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATCTATACATTAAGAGACTCGGAAATCCACATATCAAATATCCGAGTGTAGTGTAAACATTC 840

Qy       1535 CAAAAACCGTGATGGAATGGAACAACACACTTAAAAATCGCAGTATCCG 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAAAAACCGTGATGGAATGGAACAACACACTTAAAAATCGCAGTATCCG 885
```

```
Qy        700 AGACCTTTCGGTACTTCGTCCACAAACACAACTCCTCCGGCGGCGAACTTTTTCGGCGGTTGTT 759
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 AGACCTTTCGGTACTTCGTCGTCCACAAACACAACTCCTCCGGCGGCGAATTTTTTCGGCGGTGTT 64

Qy        760 ACTTGACTGGGCGACGTAATCCACGATCTCTTTTCCGTCATCGTCTTTCCGTGCTCCAAA 819
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ACTTGACTGGGCGACGTAATCCACGATCTCTTTTTCCGTCATCGTCTTTGCTCCAAA 124

Qy        820 ACAACAACGACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGCGGAACACCTGCG 879
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ACAACAACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGCGGAAGACCTGCCACGCGCCGCGCG 184

Qy        880 TCGAACAGTGTTGGGGTGTTGGAGCAAGAGATGGAATTCCAAATTCAGCGTGAGCCACCTGATAG 939
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TCGAACAGTGTGGGGTGTTGTTGTAACAATATGGAATTCCAAATTCAGCGGGCCACCTGATAT 244

Qy        940 CCTTTGTACTTAATCAGAGACTTCAGGCGGCGGTCAACGATGAAGATGTTGTTCGTTCTTCGTGCC 999
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CCTTTGTATTTAATTAAAGACTTCAAGCGGTCAACTATGAAGAAGTGTTCGTTCTTCGTGCC 304

Qy       1000 CAGTAAGCTATGTCTCCAGAAATGTAGCCATCCATCCTTCCTTGTCAATCAAGGCGTTGTCGGCT 1059
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 CAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGGCT 364

Qy       1060 TCCCGGATTGTTTACATAACCGGACATAATCATTAGGACCTCTCACACACAGTTCGCCTCTTT 1119
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TCCCGGATTGTTTACATAACCGGACATAATCATTAGGGTCCTCTCACACACAGTTCGCCTCTC 424

Qy       1120 TGATTAAACGCCCAGCGGTTTTCCCGGTATCCAGATCCAACACCTTGCTTCAAAAAAATGGA 1179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TGATTAACGCCCAGCGGTTTTCCCGGTATCCACAGATCCAACACCTTCGCTTCAAAAAAATGGA 484

Qy       1180 ACAACTTTACCGACACCGCCGCCGGGTTTATCATCCCCCTCGGGTGTTAATCAGAATAGCTGAT 1239
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 ACAACTTTACCGACACCGCCGCCGGGTTTATCATCCCCCTGGGTGTTAATCAGAATAGCTGAT 544

Qy       1240 GTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTGGCGAACCTCTTGGCAACC 1299
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GTAGTCTCAGTGAGCCCATATCCTTGTCTGTATCCCTGGAGATGGAAGGCGTTTTGCAACC 604

Qy       1300 GCTTCCCCGACTTCCTTAGAGAGGGGAGCGCCACCAGAGAGCCAATTTCGTGTGTAAATTAGAT 1359
              |||| || ||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 ...
```

```
RESULT 11
BZ679896/c
LOCUS       BZ679896                 916 bp    DNA     linear   GSS 05-FEB-2003
DEFINITION  PUBEV13TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa039C02,
            genomic survey sequence.
ACCESSION   BZ679896
VERSION     BZ679896.1  GI:28234344
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 916)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..916
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa039C02"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     Cot selected genomic DNA library"
ORIGIN

Query Match             29.0%;  Score 803.4;  DB 8;  Length 916;
  Best Local Similarity   93.2%;  Pred. No. 9.3e-154;
  Matches  851;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy       1026 CCAACCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTGTTACACCGACAT 1085
              |||| ||||||||||||| ||||| ||||| |||| ||||||||||| |||| ||||||
Db        916 CCATCCATCCTTGTCAAATCAAGGGCGTTGGGTCGCGTTCGCTTCCGGATTGTTCATAACCGGACAT 857

Qy       1086 AATCATAGGACCTCTCACACACAGTTCGCTCTTCTTTGATTAACGCCACCAGGGGTTTTTCCCGGT 1145
              |||||| || |||||| ||||| ||||||||| ||  ||||||||||||    |||| |||||||
Db        856 AATCATTAGGTCCTCTGACACATAAATTCGCTCTCTCGATTAACGCCACGCGTTTTCCCGGT 797

Qy       1146 ATCCAGATCCACAACCTTCGGCTTCAAAAAAATGGAACAACTTACCGGACCGGCGCCCCGGGTTT 1205
              ||||||||||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||
Db        796 ATCCAGATCCACAACCTTCGGCTTCAAAAAATGGAACAACTTTACCGGACCGGCGCCCCGGGTTT 737
```

```
RESULT 12
BZ682015/c
LOCUS       BZ682015                 866 bp    DNA     linear   GSS 05-FEB-2003
DEFINITION  PUBJP02TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa070A04,
            genomic survey sequence.
ACCESSION   BZ682015
VERSION     BZ682015.1  GI:28237996
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 866)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)

Qy       1206 ATCATCCCCCTCGGTGTGTAATCAGAATAGTTCAGTGATGTAGTGTCTCAGTGAGCCCATATCCTTG 1265
              ||||||||| |||||||||||||||||||| |     |||||||||| |||||||||||||||||||
Db        736 ATCATCCCCCTCGGGTGTGTAATCAGAATAGCTGAATAGCTGATGTAGTGTCAGTGAGCCCATATCCTTG 677

Qy       1266 CCTGATACCTGGCAGATGGGAACCTCTTGGCGAACCGCTTCCCGACTTCCCTTAGAGATGGGGG 1325
              |  |||||| || || ||| || ||||| |||| ||||||||||| ||| ||||||||| | |||
Db        676 TCGTATCCCCTGGGAAGGATGGGAAGCGTTTCGGAACCGGCTTCCCCGGACTTCTTTCGAAAGAGG 617

Qy       1326 AGCCGCCACCAGCAGCGAAGCAAATTTCGTGTTAAATTAGATAGAATAGAAATCGTATTTGTCAANTCAGAGTGCT 1385
              |||| ||||||||| |||||||||| ||||| |||||||||||| |||||| |||||||||||||| || ||||||
Db        616 TGCGGCCCCAGGAAGCAATTTCGTGTAAATTGATAAATCGTATTTGTCAATCAGAGTGCT 557

Qy       1386 TTTGGCGAAGAAGGAGAATAGGGGTTGGCCACCAGCAGCGAAGCAATTCGTGTAATCCTG 1445
              |||||||||| ||||| |||||||||||||||| |  ||||||||||||||||| ||||||
Db        556 TTTGGCGAAGAATGAAAAATAGGGGTTGGCTACTAGCGAACGCACTTTGAATTTTGTAATCCTG 497

Qy       1446 AAGGGCTCCTCAGAAAACAGCTCTTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACA 1505
              ||||| ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        496 AAGGGATCGTAAAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACA 437

Qy       1506 TATCAAATATCCGAGTGTAGTAAAACATTCCAAAATCTATACATTAAGACGACTCGAAAATCACA 1565
              |||||||||||||||||||||||| ||||||||||  || ||||||||||||||||||||||| |
Db        436 TATCAAATATCCGAGTGTAGTAAAACATTCCAAAATCGCACAGTATCCGAGTGTAGTAAAACACTTAA 377

Qy       1566 AAATCGCAGTATCCGGAATGGAATGATTGATTGCCAAAAATAGGATCTCTGGCCATGGCAGAATCT 1625
              |||||||||| ||||||||||||| ||| |||||||||||| |||||| |||||| |||||| || |
Db        376 AATCGCAGTATCCGGAATGGAATGATTGATTGCCAAAAATAGGATCTCTGGGCCATGGCGAGAATCT 317

Qy       1626 CACGCAGGCAGCAGTTCTATGAGGCAGAGCAGCGACACCTTTAGGCAGCAGCCAGTAGATCCAGAGGA 1685
              || |||||||| |||||||||||||||||| ||||||||| |||| ||  |||||| ||||||||||
Db        316 GACGCGGCGCAGTTCTATGCGGGAAGGGCCACACCCTTAGGTAACCCAGTAGATCCAGAGGA 257

Qy       1686 GTTCATGATCAGTGCAATTGTTCTTGTCTCCCTATCGAAAGGAAAAGGACTCTGCCACAAAATCTATTC 1745
              |||||||||||||||||||||||||||||| |||||  ||| ||||||||| |||||||| ||||| ||
Db        256 ATTCATTATCAGTGCAATTGTTTTGTTCACGGTACAAAGGACTCGGTGTACAAAAATCGTTATTC 197

Qy       1746 ATTAAAAACCGGCAGGGAGGTAGTAGATGTGAGGATGTGACGAAACGTGTACATGCAAATCCCCTGGTA 1805
              |||||||||||| |||||||| |||| ||||||||||||| ||||||| ||||||||||||||||||||
Db        196 ATTAAAAACCGGCAGGGAGGTAGTAGATGTGAGGATGTGACGAAACGTGTCATCAGCGAAAATCCCTGGTA 137

Qy       1806 ATCCCGTTTAGAATCCATGATAATAATATTTTTTGGGATGATTGGGAGGCTTTTTTTGGCACGTT 1865
              ||||||||||| ||| |||| |||||||||||||||| ||| ||||||||||||||||||||||
Db        136 ATCCCGTTTAGAATCCATGATAATAATATTTTCTGGATTATTGGGTGATTTTTTTTTGCACGTT 77

Qy       1866 CAAAAATTTTTTGCAACCCCTTTTTTGGAAACGAACACCACGGTAGGCGTGCGAAAATCCCAT 1925
              |||||||||||||||||||| ||||||||||||||||| |  ||||||||||||||||||||
Db         76 CAAAAATTTTTTGCAACCCC-TTTTCGGAAAACAAAACACTACGGTAGGCGTGCGAAAATGTTCAT 18

Qy       1926 ACTGTTGAGCAAT 1938
              |||||||||||||
Db         17 ACTGTTGAGCAAT 5
```

COMMENT   TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
source   1..866
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /clone="ZMMBTa0070A04"
            /db_xref="taxon:4577"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
            cot selected genomic DNA library"
ORIGIN
Query Match   28.7%; Score 795.6; DB 8; Length 866;
Best Local Similarity   94.9%; Pred. No. 3.6e-152;
Matches 822; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy   1395 GAAGGAGAGATAGGGGTTGGGCACCCAGCAGCGCCACCTTTGAATCTTGTTAATCCTGAAGGCTCCT 1454
          |||| ||||| |||||||||| ||||||||||| |||||||||||| |||||||||||||||| | |
Db   866 GAATGAAAAATAGGGGTTGGTACTAGCAACGCACTTTGAATTTGTTAATCCTGAAGGGGATCG 807

Qy   1455 CAGAAACAGCTCTTCTTCAAATCTATACATTATACATTAAGACGACTCGAAATCCACATATCAAATA 1514
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   806 TAAAAACAGCTCTTCTTCAAATCTATACATTATACATTAAGACGACTCGAAATCCACATATCAAATA 747

Qy   1515 TCCGAGTGTAGTAAACATTCCAAAAACCGTGATGATGGAATGGAACAACACTTAAAAATCGGCAGT 1574
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   746 TCCGAGTGTAGTAAACATTCCAAAAACCGTGATGATGGAATGGAACAACACTTAAAAATCGGCAGT 687

Qy   1575 ATCCGGAATGATTTGATTTGCCAAAAATAGGATCTCTGGCATGCCAGGAATCTCACCGCAGGC 1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   686 ATCCGGAATGATTTGATTTGCCAAAAATAGGATCTCTGGCATGCCAGGAATCTCACCGCAGGC 627

Qy   1635 AGTTCTATGGAGGCAGAGGCGACCACCTTTAGGCGACGAGCCAGTAGACGTGACGAGGAGTTCATGGAT 1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   626 AGTTCTATGCGGAGGGGCCACACCCTTAGGTAACCCAGTAGATCCAGAGGAAATTCATTAT 567

Qy   1695 CAGTGCAATTGTCTTGTCCCTATCGAAGGACTCGCGACAAAAATCGTATTCATTAAAAACC 1754
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   566 CAGTGCAATTGTTTTGTCACGGACTCGGTACAAAAGATCGTATTCATTAAAAACC 507

Qy   1755 GGGGAGGTAGATGAGATGTGACGAACGTGTGACATCGCATCGCAGAAATCCCTGGTAAATCCGTTTT 1814
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   506 GGGGAGGTAGATGAGATGTGACGAACGTGTGACATCGCATCGCAGAAATCCCTGGTAAATCCGTTTT 447

Qy   1815 AGAATCCATGATAATAATAATTTTTTGGATGATTGGGGAGCTTTTTTTGCACGTTCAAAATTTT 1874
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   446 AGAATCCATGATAATAATAATTTTCTGGATTATTGGTAATTTTTTGCACGTTCAAAATTTT 387

Qy   1875 TTGCAACCCCTTTTTGGAAACGAACACACCGGTAGGGCTCGAAATGCCCATACTGTTGAG 1934
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   386 TTGCAACCCCTTTTTGGAAACAACTACAGGTAGGCTGCGGAAAGTTCATACTGTTGAG 327

Qy   1935 CAATTCACGTTCATTATATAAAATGTGTTCGCGGGGCGGGCAACTGCAACTCCGATAAATAACGC 1994
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326 CAATTCACGTTCATTATATAAAATGTCGTTCCCGGGGCGGCAACTGCAACTCCGATAAATAACGC 267

Qy   1995 GCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATGACGAGGACGATTCTGTGATT 2054
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266 GCCCAACACCGGCATAAAGAATTGAAAGAGAGTTTTCACTGCATAACGACCGGATTCTGTGATT 207

Qy   2055 TGTATTCAGCCCATATCGTTTCGCTTCGCCAACCGAACCGACCAATTCTGAAGTGACTC 2114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 TGTATTCAGCCCATATCGTTTCATTAGCTTCTGCCAACCGAACGGACACATTCGAAGTATTC 147

Qy   2115 AGCCTAAGTAGTGATGTGTCCACCTCGATATGTCAACTCTGTAAAAGCAATTGTTCCAGGAACCAG 2174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   147 AGCCTAAGTAGTGATGTGTCCACCTCGATATGTCAACTCTGTAAAAGCAATTGTTCCAGGAACCAG 2174

RESULT 13

LOCUS   BZ677942/c
Db   146   BZ677942   PUBFU55TD ZM_0.6_1.0_KB   892 bp   DNA   linear   GSS 05-FEB-2003
DEFINITION PUBFU55TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa045J13,
            genomic survey sequence.
ACCESSION   BZ677942
VERSION   BZ677942.1   GI:28230797
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 892)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE   Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
source   1..892
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa045J13"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
            cot selected genomic DNA library"
ORIGIN
Query Match   28.7%; Score 795.2; DB 8; Length 892;
Best Local Similarity   95.0%; Pred. No. 4.4e-152;
Matches 821; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Db   146 CCGGTAGGTGAGTGTCACCTCGATAATGTGCAATCTGTAAAAGCAATTGTTCCAGGACGGATA 87
          |||| |||||||||| |||||||||||||||||||| ||||||||||||||| |||||||||||
Qy   2175 GGCGTATCTCTTCATTAGCCTTATGCAGTTGCTCTTCCAGCCGGTTCCATCTTCCAGCGCGGATA 2234

Db   86 GGCGTATCTCTTCATTAGCCTTATGCAGTTGCTCTTCCAGCGCGGTTCCATCCTCTAGAGGATA 27
          ||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||||||||
Qy   2235 GAATGGGCGCCCGGCCCGCCCTTTCTTTTATGT 2260

Db   26 GAATGGGCGCCCGGCCCGCCCTTTCTTTTATGT 1

Db   146 GCAGCGCACTTTGAATCTTGTAATCCTGAAGGCTCCTCAGAAAACAGCTCTTCTTCAAATC 1477
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1418 GCAGCGCACTTTGAATCTTGTAATCCTGAAGGCTCCTCAGAAAACAGCTCTTCTTCAAATC 1477

Db   892 GCAACGCACTTTGAATTTGTAATCCTGAAGGGGATCGTAAAAACAGCTCTTCTTCAAATC 833
          |||||||||||||||| |||||||||||||||| |||||||||||||||||||||||||||
Qy   1478 TATACATTATAAGACGACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAACATTCCAA 1537

Db   832 TATACATTAAGACGACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAACATTCCAA 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1538 AACCGTGATGATGGAATGGAACAACACTTAAAAATCGGCAGTATCCGGAATGATTTGATTTGCCAA 1597

Db   772 AACCGTGATGGAATGGAACAACACTTAAAAATCGGCAGTATCCGGAATGATTTGATTTGCCAA 713
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1598 AAATAGGAGATCTCTGGCATCGGAGAATCTCACGCAGGCAGTTCTATGGAGGCAGAGGCGACGAGC 1657

Db   712 AAATAGGAGATCTCTCGGCATCGGAGAATCTGACGCAGGCAGTTCTATGCGGAAGGGCCACAC 653
          ||||||||||||||| ||||||||||||||| |||||||||||||||| ||||||||||||||
Qy   1658 CTTTAGGCGAGACCAGTAGATCCAGAGGAGTTCATGATCAGTGCAATTGTGTCTTGTCCCTAT 1717

```
RESULT 14
BZ682835
LOCUS       BZ682835               882 bp    DNA     linear   GSS 05-FEB-2003
DEFINITION  PUBGP67TD ZM_0_6_1.0_KB zea mays genomic survey sequence.
ACCESSION   BZ682835
VERSION     BZ682835.1  GI:28239565
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     TIGR
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..882
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa050L14"
                     /clone_lib="ZM_0_6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

Query Match    28.6%;  Score 793.4;  DB 8;  Length 882;
Best Local Similarity   94.7%;  Pred. No. 1e-151;
Matches  832;  Conservative    0;  Mismatches   46;  Indels    1;  Gaps    1;
```

```
Qy    652 CCTTAGGTAACCCAGTAGATCCAGAGAGAATTCATTATCAGTGCAATTGTTTTGTCACGAT 593
Db   1718 CGAAGGACTCTGGCACAAAAATCGTATTCATTAAAACCGGGAGGTAGAGTGACGA 1777

Qy    592 CAAAGGACTCTGGTACAAAAATCGTATTCATTAAAACCGGGAGGTAGAGTGACGA 533
Db   1778 ACGTGTACATCGACTGAAAATCCCTGGTAATCCGTTTTTAGAATCCATGATAATAATTTTT 1837

Qy    532 ACGTGTACATCGACTGAAAATCCCTGGTAATCCGTTTTTAGAATCCATGATAATAATTTTCT 473
Db   1838 GGATGATTGGGGAGCTTTTTTCCACGTTCAAAATTTTTTGCAACCCCCTTTTTTGGAAACGA 1897

Qy    472 GGATTATTGGGTAATTTTTTTTGCACGTTCAAAATTTTTTGGAACCCCTTTTTTGGAAACAA 413
Db   1898 ACACCACGGTAGGCTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTATATAAATGT 1957

Qy    412 ACACTACGGTAGGCTGCGAAATGTTCATACTGTTGAGCGAATTCACGTTCATTATATAAATGT 353
Db   1958 CGTTCGGGGGCGGCGCCAACCTGCCAACACCGGGCCCCAACACCGCCATAAAAGAATT 2017

Qy    352 CGTTCGGGGGCGGCGCCAACCTGCCAACACCGGGCCCCAACACCGCCATAAAAGAATT 293
Db   2018 GAAGAGAGTTTTCACTGCATACGACGATTCTGTGATTTGTATTCAGCCCCATATCGTTTCA 2077

Qy    292 GAAGAGAGTTTTCACTGCATACGACGATTCTGTGATTGTTGTATTCAGCCCATATCGTTTCA 233
Db   2078 TAGCTTCTGCCAACCGAACGGACATTTGTGAAGTACTCAGCGTGAGTGATGTCCACCTCGA 2137

Qy    232 TAGCTTCTGCCAACCGAACGGACATTTGTGAAGTATTCGGTACGTGATGTTCACCTCGA 173
Db   2138 TATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCACCAGGGGTATCTCTTCATGAGCCTTAT 2197

Qy    172 TATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCACCAGGGGTATCTCTTCATGAGCCTTAT 113
Db   2198 GCAGTTGCTCTCCAGCCGGTTCCATCTTCACGACCGGATAGAATGGCGCCGCGGCCTTCTTA 2257

Qy    112 GCAGTTGCTCTCCAGCCGGTTCCATCTTCACGACCGGATAGAATGGCGCCGCGGCCTTCTTA 53
Db   2258 TGTTTTTGGCGGTCTTTCCATGGGAC 2281

Qy     52 TGTTTTTGGCGGTCTTTCCATTTTAC 29
```

```
Qy    700 AGACCTTTCGGGTACTTCGTCCACAACAACTCCTCCGGCCAACTTTTTCGGCGGTTGTT 759
Db      5 AGACCTTTCGGGTACTTCGTCCACAAAACAACTCCTCCGGCCAACTTTTTTCGGCGGTTGTT 64

Qy    760 ACTTGACTGGCGACGGTAATCCACGGATCTCTCTTTTTTCCGTCATCGTCTTTCCGGTCCAAA 819
Db     65 ACTTGACTCGGCGACGGTAATCCACGGATCTCTCTTTTTTCCGTCATCGTCTTTCCGGTCCAAA 124

Qy    820 ACACACAAGGCGCGGGGAGAGTTCACCGCCGTCAATCGTGGGGAGAGACCCTGCGGACACTGCG 879
Db    125 ACAACAACGGCGGGGGGAAGAGTTCAACCGCCGTCATCGTCGGGAAGACCCTCCACGGCGCCGCG 184

Qy    880 TCGAAGAGTTGGGGGTGTTGGAGCAAGAGGATGGATTCCAAATTCAGCGGGGCCACACTGATAG 939
Db    185 TCGAAGAGTGGGGGTGTTGTAACAAGATCATTCCAAATTCAGCGGGGCCCACCTGATAT 244

Qy    940 CCTTTGTACTTAATCAGAGAGACTTCAGCGCCGTCAACGATGAAGAAGAAGAGTGTCGTCTCC 999
Db    245 CCTTTGTAGTTAATTAAAAGACTTCAAGCGGGTCAACTATGAAAGTGTTTCGTCTTCGTCC 304

Qy   1000 CACGTAAGCTATGTCTCCAGAAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCT 1059
Db    305 CAGTAAGCTATGTCTCCAGAAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCT 364

Qy   1060 TCCGGATTGTTTACATAACCGGACATAATCATAGGACCTCTCACACACAGTTCGCCTCTTT 1119
Db    365 TCCGGATTGTTTACATAACCGGACATAATCATAGGTCCTCTCACACACAGTTCGCCTCTC 424

Qy   1120 TGATTAACGCCCAGCGTTTTCCCGGTATCCAGATCCACAACCTTCGCCTTCAAAAAAATGGA 1179
Db    425 TGATTAACGCCCAGCGTTTTCCCGGTATCCAGATCCACAACCTTCGCCTTCAAAAAATGGA 484

Qy   1180 ACAACTTTACCGACCGGCGCCGGCCCCCGGTTTATCATTCCCCTCGGGGTGTAAATCAGCTGAT 1239
Db    485 ACAACTTTACCGACCGGCGCCGGCCCCCGGTTTATCATCCCCTCGGGTGTAAATAGCCTGAT 544

Qy   1240 GTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTGGCAAGAGGAGGAACCTTCTTGGCCACC 1299
Db    545 GTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTGGCAAGAGGAGGAACCTTCTTGGCCACC 604

Qy   1300 GCTTCCCCGCGACTTCCTTAGAGAGGAGCGAGCGGCCACCCAGGAAGCAAATTTCGTGTAAATTAGAT 1359
Db    605 GCTTCCCCGGACTTCTTTCGAAAGGCGCCCCCCAGAAGCAATTTCGTGTAAAAATTAGAT 664

Qy   1360 AAATCGTATTTGTCAAATCCAGAGTGGCTTTTGGCAAGAAGAGGAGAATAGGGTTGGCCACCAGC 1419
Db    665 AAATCGTATTTGTCAAATCAGAGTGCTTTTGGCGAAGAATGAAAATAAGGGTTGGGTTACTAGC 724

Qy   1420 AGGCACACTTTGAATCTTGTAATCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAAATCTA 1479
Db    725 AACGGCACTTTGAATTTGTAATCCTGAAGGGATCGTAAAAACAGCTCTTCTTCAAAATCTA 784

Qy   1480 TACATTAAGACGACTCGAAATCCACATATCAAATATCCGAGTGTTAGTAAAAACATTCCAAA 1539
Db    785 TACATTAAGACGACTCGAAATCCACATATCAAATATCCGAGTGTTAGTAAAACATTCCAAA 844

Qy   1540 CCGTGATCGGAATGGAACACACTTAAAATCGCAGTATCC 1578
Db    845 CCGTGATCGGAAATGG-ACAACACTTAAAATCGCAGTATCC 882
```

```
RESULT 15
BZ683904/c
LOCUS       BZ683904              867 bp    DNA     linear   GSS 05-FEB-2003
DEFINITION  PUBHO3TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa05?I09,
            genomic survey sequence.
ACCESSION   BZ683904
VERSION     BZ683904.1  GI:28241625
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 867)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..867
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa05?I09"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     Cot selected genomic DNA library"
ORIGIN
```

Query Match 28.4%; Score 787.6; DB 8; Length 867;
Best Local Similarity 95.4%; Pred. No. 1.6e-150;
Matches 811; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy    1427 TTTGAATCTTGTAATCCTGAAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTA 1486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     447 GTAATTTTTTTGCAGTTCAAAATTTTTGCAACCCCTTTTGGAAAACAAACACTACGG 388

Qy    1487 AGACGACTCGAAATCCACATATCAAATATCCGACTGTAGTAAACATTCAAAAACCGTGAT 1546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     807 TTTGAATTTTGTAATCCTGAAGGGATCGTAAAAACAGCTCTTCTTCAAATCTATACATTA 808

Qy    1547 GGAATGGGAACAACACTTAAAAATGCCAGTATCCGGAATGATTTGATTGCCAAAAAATAGGAT 1606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     747 GGAATGGAACAACACTTAAAAATGGCAGTGATCCGGAATGATTTGATTGCCAAAAAATAGGAT 688

Qy    1607 CTCTGGGCATGCGAGAATCTCACGCAGGCAGAGGCGACACCTTTAGGCA 1666
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 CTCCGGCCATGCGAGAATCTGACGCAGGCGACGTTCTCATGCCGGAAGGGCCCACACCCTTAGGTA 628

Qy    1667 GACCAGTAGATCCAGAGGAGTTCATGACTCAGTGCAAATTGTCTCTTGTGTCCCTTATCGAAGGACT 1726
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     627 ACCCAGTAGATCCAGAGAATCCATTAATCAGTGCAAATTGTTTGTTCTCACCGATCAAAGGACT 568

Qy    1727 CTGGCACAAAATCTATTCATTAAAAACCCGGGAGGTAGATTGAGATGTGACGAACGTGTACA 1786
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     567 CTGCTACAAAATGTATTCATTAAAAACCGGGAGGTGAGATTGAGATGTGACGAACGTGTACA 508

Qy    1787 TCGACTGAAATCCCTGGTAATCCGTTTAGAATGAATAAATAATTTTTTGGATGATTG 1846
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     507 TCGACTGAAAATCCCTGGTAATCCGTTTAGAATCATCATGATAATAATTTTTCTGGATTATTG 448

Qy    1847 GGAGCTTTTTTTTGCACGTTCAAAATTTTTTGCAACCCCTTTTGGAAAACGAACACCACGG 1906
           || ||||||||||||||||||||||||||||||||||||||||||||| |||
Db
```

```
Db     447 GTAATTTTTTTGCAGTTCAAAATTTTTGCAACCCCTTTTGGAAAACAAACACTACGG 388
Qy    1907 TAGGCTGCGAAATGCCCATACTGTTGAGCAATTCACGTTCATTATTAAAAATGTCGTTCGCGG 1966
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     387 TAGGCTGCGGAAATGTTCATTACTGTTGAGCAATTCACGTTCATTATTAAAAATGTCGTTCGCGG 328
Qy    1967 GCGCAACTGCAACTCCGATAAATAACGCGCCCAACCACCCGGCCATAAAAGAAATTGAGAGAGT 2026
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327 GCGCAACTGCAACTCCGATAAATAACGCGCCCAACCACCCGGCCATAAAAGAAATTGAGAGAGT 268
Qy    2027 TTTCACTGCATACGACGATTCGTGATTTGTTATTCAGCCCATATCGTTTCATTAGCTTCTG 2086
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267 TTTCACTGCATACGACGATTCGTGATTTGTTATTCAGCCCATATCGTTTCATTAGCTTCTG 208
Qy    2087 CCAACCGGAACGGACATTTCGAAGTACTCACGTAAGTGAAGTGATGTTCAACCTCGATATGTGCAT 2146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     207 CCAACCGGAACGGACATTTCGAAGTATTCCGCGTAGTGAAGTGATGTTCAACCTCGATATGTGCAT 148
Qy    2147 CTGTAAAAGCAAATTGTTCTTCCAGGGAACCAGAGGCGCGTATCTTCTTCATTAGCCTTATGCAGTTGCT 2206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147 CTGTAAAAGCAAATTGTTCCAGGGAACCAGAGGCGCGTATCTTCTTCATTAGCCTTATGCAGTTGCT 88
Qy    2207 CTCCACGCGGTTCCATCTTCCAGCCGGATAGAATGGGCGCCCGGGCCCTTTCTTTAGTTTTTTGG 2266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87 CTCCACGGCGGTTCCATCCTCTAGAGGAATAGAATGGGCGCCCGGGCCTTTCTTTAGTGTTTTTTGG 28
Qy    2267 CGTCTTCCAT 2276
           ||||||||||
Db      27 CGTCTTCCAT 18
```

Search completed: September 29, 2005, 17:34:28
Job time : 8771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2005, 11:18:45 ; Search time 454 Seconds
                 (without alignments)
                 9987.055 Million cell updates/sec

Title: US-10-734-801-18
Perfect score: 2771
Sequence: 1 ggatccgctgtggaatgtgt..........atgggaccacaaatctctc 2771
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents NA:*
                 1:/cgn2_6/ptcdata/1/ina/5A_COMB.seq:*
                 2:/cgn2_6/ptcdata/1/ina/5B_COMB.seq:*
                 3:/cgn2_6/ptcdata/1/ina/6A_COMB.seq:*
                 4:/cgn2_6/ptcdata/1/ina/6B_COMB.seq:*
                 5:/cgn2_6/ptcdata/1/ina/PCTUS_COMB.seq:*
                 6:/cgn2_6/ptcdata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2771	100.0	2771	4	US-10-066-130-18	Sequence 18, Appl
2	2771	100.0	2771	4	US-10-066-130-17	Sequence 17, Appl
3	2674	96.5	2674	4	US-10-066-130-19	Sequence 19, Appl
4	2327	84.0	2327	4	US-10-066-130-20	Sequence 20, Appl
5	1656.4	59.8	13554	4	US-09-743-194-28	Sequence 28, Appl
6	1655.6	59.7	1686	4	US-09-602-628-7	Sequence 7, Appl
7	1655	59.7	12850	4	US-09-743-194-22	Sequence 22, Appl
8	1654.2	59.7	14194	4	US-09-577-424-3	Sequence 3, Appl
9	1653.4	59.7	12614	4	US-09-577-424-1	Sequence 1, Appl
10	1651.4	59.6	13073	4	US-09-743-194-19	Sequence 19, Appl
11	1651.4	59.6	12844	4	US-09-743-194-20	Sequence 20, Appl
12	1649	59.5	1650	1	US-08-354-240A-3	Sequence 3, Appl
13	1647.6	59.5	1686	1	US-08-354-240A-5	Sequence 5, Appl
14	1643	59.3	1653	4	US-09-602-628-1	Sequence 1, Appl
15	1642.8	59.3	1686	4	US-09-602-628-1	Sequence 1, Appl
16	1640.2	59.2	12851	4	US-09-743-194-21	Sequence 21, Appl
17	1626	58.7	1685	4	US-09-602-628-5	Sequence 5, Appl
18	1560	56.3	14113	4	US-09-223-134-1	Sequence 1, Appl
19	1560	56.3	14113	4	US-09-223-135-1	Sequence 1, Appl
20	1560	56.3	14113	4	US-08-992-801-1	Sequence 1, Appli
21	1520.6	54.9	14413	3	US-09-223-535-1	Sequence 1, Appli
22	1519	54.8	5789	3	US-09-282-996-1	Sequence 1, Appli
23	1519	54.8	5791	3	US-08-862-431-32	Sequence 32, Appl
24	1519	54.8	5793	3	US-08-862-431-31	Sequence 31, Appl
25	1519	54.8	5793	3	US-08-862-431-30	Sequence 30, Appl
26	1518.4	54.8	5427	3	US-09-282-996-2	Sequence 2, Appli
27	1518.4	54.8	10160	3	US-09-097-319A-8	Sequence 8, Appl
28	1518.4	54.8	10160	4	US-09-643-971-8	Sequence 8, Appli
29	1518.4	54.8	11784	3	US-09-097-319A-9	Sequence 9, Appli
30	1518.4	54.8	11784	3	US-09-643-971-9	Sequence 9, Appli
31	1518.4	54.8	11991	3	US-09-097-319A-10	Sequence 10, Appl
32	1518.2	54.8	11991	3	US-09-643-971-10	Sequence 10, Appl
33	1518.2	54.8	1811	3	US-08-867-352-22	Sequence 22, Appli
34	1518.2	54.8	1818	2	US-08-536-559A-3	Sequence 3, Appli
35	1518.2	54.8	5819	3	US-08-536-559A-2	Sequence 2, Appli
36	1518.2	54.8	5819	3	US-08-862-431-27	Sequence 27, Appl
37	1518.2	54.8	5819	3	US-08-536-559A-4	Sequence 4, Appli
38	1518.2	54.8	5938	2	US-08-862-431-28	Sequence 28, Appl
39	1518.2	54.8	6092	3	US-08-536-559A-1	Sequence 1, Appli
40	1518.2	54.8	6092	3	US-08-862-431-26	Sequence 26, Appl
41	1518.2	54.8	8937	4	US-09-872-733A-8	Sequence 8, Appli
42	1518.2	54.8	8937	4	US-09-872-733A-9	Sequence 9, Appli
43	1518.2	54.8	11616	1	US-08-196-259-2	Sequence 2, Appli
44	1518	54.8	1722	3	US-08-380-061B-1	Sequence 18, Appli
45	1516.8	54.7	6044	1	US-08-316-950-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-066-130-18

; Sequence 18, Application US/10066130
; Patent No. 6699657
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: In Vitro System for Replication of RNA-Dependent RNA Polymerase (
; TITLE OF INVENTION: Viruses
; FILE REFERENCE: PH-7171 NP
; CURRENT APPLICATION NUMBER: US/10/066,130
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/265,437
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: viral
US-10-066-130-18

Query Match   100.0%; Score 2771; DB 4; Length 2771;
Best Local Similarity 100.0%;
Matches 2771; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGATCCGCTGTGGAATGTGTGTCGTCAGTTAGGGTGTGGGAAAGTCCCCAGGCTCCCCAGCAGG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGATCCGCTGTGGAATGTGTGTCAGTTAGGGTGTGGGAAAGTCCCCAGGCTCCCCAGCAGG 60

Qy   61 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120

Qy   121 CTCCCCAGCAGCGAGAAGTATGCAAAGCATCTCAATTAGTCAGCAGCAACCATAGTCCC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTCCCCAGCAGCGAGAAGTATGCAAAGCATCTCAATTAGTCAGCAACCATAGTCCC 180

Qy   181 GCCCCTAACTCCGCGCCCATCCCGCGCCCCTAACTCCGCGCCCAGTTCCGCGCCCCA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GCCCCTAACTCCGCGCCCATCCCGCGCCCCTAACTCCGCGCCCAGTTCCGCGCCCCA 240

Qy   241 TGGCTGACTAATTTTTTTTTATTTTATTTTATGCAGAGGCCGAGGCCGAGGCCTATT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGCTGACTAATTTTTTTTTATTTTATTTTATGCAGAGGCCGAGGCCGAGGCTATT 300

Qy   301 CCAGAAGTAGTGAGGAGGCCTTTTTTTGGAGGCCTTAGCCTTTTGCAAAAAGCTTACATGATC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCTTAGCCTTTTGCAAAAAGCTTACATGATC 360

Qy   361 TGCAGAGAGGCCAGTATCAGCACTCTCTCGAGTCATGCGGCTCACGGACCTTTCACAGCT 420

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGCAGAGAGGCCAGTATCAGCACTCTCTGCAGTCATGCGGCTCACGGACCTTTCACAGCT 420

Qy         421 AGCCGTGACTAGGGCTAAGATGGAGCCACCATTAAAGAAGGAAGGAAAAGAAAGGAAAAA 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 AGCCGTGACTAGGGCTAAGATGGAGCCACCATTAAAGAAGGAAGGAAAAGAAAGGAAAAA 480

Qy         481 AGAAGGAAAGAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAGAAAAAAAAAA 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGAAGGAAAGAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAGAAAAAAAAAA 540

Qy         541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAATGGCCTAAGA 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAATGGCCTAAGA 600

Qy         601 GGCCGGAGTGTTTACCCCAACCTTTAAACGGCGATCTTTCCGCCCTTCTTGGCCTTTATG 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GGCCGGAGTGTTTACCCCAACCTTTAAACGGCGATCTTTCCGCCCTTCTTGGCCTTTATG 660

Qy         661 AGGATCTCTCTGATTTTTCTTGCGTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCC 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AGGATCTCTCTGATTTTTCTTGCGTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCC 720

Qy         721 ACAAACACAACTCCTCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCC 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 ACAAACACAACTCCTCCGCGCAACTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCC 780

Qy         781 ACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACAACAACGGCGGCGGGAAGT 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 ACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTCCAAAACAACAACGGCGGCGGGAAGT 840

Qy         841 TCACCGGCGTCATCGTCGGGAAGACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGG 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 TCACCGGCGTCATCGTCGGGAAGACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGG 900

Qy         901 AGCAAGATGGATTCCAATTCAGCGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGAC 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AGCAAGATGGATTCCAATTCAGCGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGAC 960

Qy         961 TTCAGGCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAA 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TTCAGGCGGTCAACGATGAAGAAGTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAA 1020

Qy        1021 TGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCG 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 TGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCG 1080

Qy        1081 GACATAATCATAGGACCTCTCACACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTC 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GACATAATCATAGGACCTCTCACACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTC 1140

Qy        1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCC 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCC 1200

Qy        1201 GGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATAT 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GGTTTATCATCCCCCTCGGGTGTAATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATAT 1260

Qy        1261 CCTTGCCTGATACCTGGCAGATGGAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAG 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 CCTTGCCTGATACCTGGCAGATGGAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAG 1320

Qy        1321 AGGGGAGCGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGA 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 AGGGGAGCGCCACCAGAAGCAATTTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGA 1380

Qy        1381 GTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAA 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 GTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAA 1440

Qy        1441 TCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAAT 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 TCCTGAAGGCTCCTCAGAAACAGCTCTTCTTCAAATCTATACATTAAGACGACTCGAAAT 1500

Qy        1501 CCACATATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 CCACATATCAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560

Qy        1561 CTTAAAATCGCAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAG 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CTTAAAATCGCAGTATCCGGAATGATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAG 1620

Qy        1621 AATCTCACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCA 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 AATCTCACGCAGGCAGTTCTATGAGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCA 1680

Qy        1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCG 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 GAGGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCG 1740

Qy        1741 TATTCATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCC 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 TATTCATTAAAACCGGGAGGTAGATGAGATGTGACGAACGTGTACATCGACTGAAATCCC 1800

Qy        1801 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGC 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 TGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGC 1860

Qy        1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATG 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATG 1920

Qy        1921 CCCATACTGTTGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACT 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 CCCATACTGTTGAGCAATTCACGTTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACT 1980

Qy        1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 CCGATAAATAACGCGCCCAACACCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040

Qy        2041 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGAC 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 ACGATTCTGTGATTTGTATTCAGCCCATATCGTTTCATAGCTTCTGCCAACCGAACGGAC 2100

Qy        2101 ATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATT 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 ATTTCGAAGTACTCAGCGTAAGTGATGTCCACCTCGATATGTGCATCTGTAAAAGCAATT 2160

Qy        2161 GTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 GTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2220

Qy        2221 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGGGA 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 TCTTCCAGCGGATAGAATGGCGCCGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGGGA 2280

Qy        2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTGCTCATGATGCAC 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 CGTCGGTTGGTGTTACGTTTGGTTTTTCTTTGAGGTTTAGGATTCGTGCTCATGATGCAC 2340

Qy        2341 GGTCTACGAGACCTCCCGGGGCACTCGCAAGCACCCTATCAGGCAGTACCACAAGGCCTT 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 GGTCTACGAGACCTCCCGGGGCACTCGCAAGCACCCTATCAGGCAGTACCACAAGGCCTT 2400

Qy        2401 TCGCGACCCAACACTACTCGGCTAGCAGTCTTGCGGGGGCACGCCCAAATCTCCAGGCAT 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 TCGCGACCCAACACTACTCGGCTAGCAGTCTTGCGGGGGCACGCCCAAATCTCCAGGCAT 2460

Qy        2461 TGAGCGGGGTTATCCAAGAAAGGACCCGGTCGTCCTGGCAATTCCGGTGTACTCACCGGT 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 TGAGCGGGGTTATCCAAGAAAGGACCCGGTCGTCCTGGCAATTCCGGTGTACTCACCGGT 2520

Qy        2521 TCCGCAGACCACTATGGCTCTCCCGGGAGGGGGGGTCCTGGAGGCTGCACGACACTCATA 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 TCCGCAGACCACTATGGCTCTCCCGGGAGGGGGGGTCCTGGAGGCTGCACGACACTCATA 2580
```

RESULT 2
US-10-066-130-17
; Sequence 17, Application US/10066130
; Patent No. 6699657
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: In Vitro System for Replication of RNA-Dependent RNA Polymerase
; TITLE OF INVENTION: Viruses
; FILE REFERENCE: PH-7171 NP
; CURRENT APPLICATION NUMBER: US/10/066,130
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/265,437
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 17
; LENGTH: 5860
; TYPE: DNA
; ORGANISM: viral
US-10-066-130-17

Query Match   100.0%;   Score 2771;   DB 4;   Length 5860;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 2771;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy      1 GGATCCGCTGTGGAAATGTGTGGTGTCAGTTAGGGGTGTGGAAAGTCCCCAGGCTCCCCAGGCAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGATCCGCTGTGGAAATGTGTGGTGTCAGTTAGGGGTGTGGAAAGTCCCCAGGCTCCCCAGGCAGG 60

Qy     61 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTGTCAGCAACACCAGGTGTGGAAAGTCCCCAGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGAAGTATGCAAAAGCATGCATCTCAATTAGTGTCAGCAACACCAGGTGTGGAAAGTCCCCAGG 120

Qy    121 CTCCCCAGCAGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCCCCAGCAGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCC 180

Qy    181 GCCCCCTAAACTCCGCCCACCGCCCCCGCCCTAAACTCCGCCCATTCTCCGCCCCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCCCCCTAAACTCCGCCCACCGCCCCCGCCCTAAACTCCGCCCATTCTCCGCCCCA 240

Qy    241 TGGGCTGACTAATTTTTTTTATTATTATGCAGAGGACCGCGGCAAGCGCTATT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGGCTGACTAATTTTTTTTATTATTATGCAGAGGACCGCGGCAAGCGCTATT 300

Qy    301 CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGCGTTTGCAAAAAGCTTACATGATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGCGTTTGCAAAAAGCTTACATGATC 360

Qy    361 TGCAGAGAGGCCAGTATCAGCACACTCTTGCAGTCATCGGCTCAGCGGGACCTTTCACAGCT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGCAGAGAGGCCAGTATCAGCACACTCTTGCAGTCATCGGCTCAGCGGGACCTTTCACAGCT 420

Qy    421 AGCCGTTGACTAGGGCTTAAGATGGAGCCCACCATTAAAGAAGGAGGAAAGAAAGGAAAAGAAAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGCCGTTGACTAGGGCTTAAGATGGAGCCCACCATTAAAGAAGGAGGAAAGAAAGGAAAAGAAAA 480

Qy    481 AGAAGGAAAAGAAAGAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAGGAAAAAAAGAAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AGAAGGAAAAGAAAGAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAGGAAAAAAAGAAAA 540

Qy    541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600

Qy    601 GGCCGGGAGTGTTTACCCCAACCTTTAAACGGGCCGATCTTTCCGCCCTTCTTGGCCTTTATG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCGGGAGTGTTTACCCCAACCTTTAAACGGGCCGATCTTTCCGCCCTTCTTGGCCTTTATG 660

Qy    661 AGGATCTCTCTGATTTTTTCTTGGGTCGAGTTTTTCCGGGTCGAGTTTTACTTCGGTCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AGGATCTCTCTGATTTTTTCTTGGGTCGAGTTTTTCCGGGTCGAGTTTTACTTCGGTCC 720

Qy    721 ACAAACACAACTCCTCCGCGCGCAACTTTTTTCCGGGGTTGTTACTTGACTGGCGAATCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACAAACACAACTCCTCCGCGCGCAACTTTTTTCCGGGGTTGTTACTTGACTGGCGAATCC 780

Qy    781 ACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTTCAAAACAACAACGGCGGCGGGGAGT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ACGATCTCTTTTTCCGTCATCGTCTTTCCGTGCTTCAAAACAACAACGGCGGCGGGGAGT 840

Qy    841 TCACCCGGCGGGTCATCGTCGGGGAACGACCTCCGACACCGGTGGTGTTGG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCACCCGGCGGGTCATCGTCGGGGAACGACCTCCGACACCGGTGGTGTTGG 900

Qy    901 AGCAAGATCGAATTCCAAТTCAGCGGGAGCCACCTTTGTACTTAATCAGGAGAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AGCAAGATCGAATTCCAAТTCAGCGGGAGCCACCTTTGTACTTAATCAGGAGAC 960

Qy    961 TTCAGGCGGTCAACGATGAAGAAGGTGTTCGTTCTTCGTCCCAGTAAGCTATGTCTCCAGA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTCAGGCGGTCAACGATGAAGAAGGTGTTCGTTCTTCGTCCCAGTAAGCTATGTCTCCAGA 1020

Qy   1021 TGTAGCCATCCATCCTTGTCAATCAAGCGGTTCGTCGCTTCCGGATTGTTTACATAAACCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTAGCCATCCATCCTTGTCAATCAAGCGGTTCGTCGCTTCCGGATTGTTTACATAAACCG 1080

Qy   1081 GACATAATCATAGGACCTCTCACACACAGTTCGCCCTCTTTGATTAAGGCCCCAGCGTTTTC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GACATAATCATAGGACCTCTCACACACAGTTCGCCCTCTTTGATTAAGGCCCCAGCGTTTTC 1140

Qy   1141 CCGGTATCCAGATCCACAACCTTCGCCTTCAAAAAATGGAACAACTTACCGACCGGCGCC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCGGTATCCAGATCCACAACCTTCGCCTTCAAAAAATGGAACAACTTACCGACCGGCGCC 1200

Qy   1201 GGTTTATCATCCCCTCGGTGTAATCAGAATAGCTGAGTGTAGTCTTCAGTGAGCCCATAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GGTTTATCATCCCCTCGGTGTAATCAGAATAGCTGAGTGTAGTCTTCAGTGAGCCCATAT 1260

Qy   1261 CCTTGCCTGATACCTGGCGATGGAACCTTCTTGGGCACACGGCCTTCCCTTAGAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CCTTGCCTGATACCTGGCGATGGAACCTTCTTGGGCACACGGCCTTCCCTTAGAG 1320

Qy   1321 AGGGGAGGGGCCACCGAGGAGCAATTTCGTGTGTAAATTAGATAAATCGTATTGTCAAATCAGA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGGGGAGGGGCCACCGAGGAGCAATTTCGTGTGTAAATTAGATAAATCGTATTGTCAAATCAGA 1380

Qy   1381 GTGCTTTTGGGCGAAGAAGGAGAATAGGGTTGGCACCAGCGCGCACTTTGAATCTTGTAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTGCTTTTGGGCGAAGAAGGAGAATAGGGTTGGCACCAGCGCGCACTTTGAATCTTGTAA 1440

Qy   1441 TCCTTGAAGGCTCCTCAGAAAACAGCTCTTCTTCAAAATCTATACATTAAGCACGACTCGAAAT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCCTTGAAGGCTCCTCAGAAAACAGCTCTTCTTCAAAATCTATACATTAAGCACGACTCGAAAT 1500

Qy   1501 CCAGATATCAAAATATCCGGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CCAGATATCAAAATATCCGGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560
```

```
Qy   2581 CTAACCCCATGGCTAGACGCGCTTTCTCGCGTGAAGACAGTAGTTCCTCACAGGGGGAGTGATT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CTAACCCCATGGCTAGACGCGCTTTCTCGCGTGAAGACAGTAGTTCCTCACAGGGGGAGTGATT 2640

Qy   2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGGCTCGGGCGGCCGGGCCCAGCCTCCTCGC 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGGCTCGGGCGGCCGGGCCCAGCCTCCTCGC 2700

Qy   2701 TGGGCGCCGGGCTGGGGCAACATTCCGAGGGGGACCCGTCCCCTCGGTAATGGGCGGAATGGGGACCC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TGGGCGCCGGGCTGGGGCAACATTCCGAGGGGGACCCGTCCCCTCGGTAATGGGCGGAATGGGGACCC 2760

Qy   2761 ACAAATCTCTC 2771
          |||||||||||
Db   2761 ACAAATCTCTC 2771
```

RESULT 3
US-10-066-130-19

```
;  Sequence 19, Application US/10066130
;  Patent No. 6699657
;  GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: In Vitro System for Replication of RNA-Dependent RNA Polymerase (f
;  TITLE OF INVENTION: Viruses
;  FILE REFERENCE: PH-7171 NP
;  CURRENT APPLICATION NUMBER: US/10/066,130
;  CURRENT FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/265,437
;  PRIOR FILING DATE: 2001-01-31
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
;  SEQ ID NO 19
;  LENGTH: 2674
;  TYPE: DNA
;  ORGANISM: viral
US-10-066-130-19

Query Match       96.5%;   Score 2674;   DB 4;   Length 2674;
Best Local Similarity  100.0%;   Pred. No. 0;   Mismatches  0;   DB
Matches 2674;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;
```

Qy   1 GGATCCGGCTGTGGAATGTGTCGTTCAGTTAGGGTGGGAAAGTCCCCAGCGCAGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGATCCGGCTGTGGAATGTGTCGTTCAGTTAGGGTGGGAAAGTCCCCAGCGCAGG 60

Qy   61 CAGAAGTATGCAAAGCAGCATCTCAAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGAAGTATGCAAAGCAGCATCTCAAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGG 120

Qy   121 CTCCCCAGCAGCAGGAGAATGCAAAGCATGCATTCAAATTAGTCAGCAACCATAGTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTCCCCAGCAGCAGGAGAATGCAAAGCATGCATTCAAATTAGTCAGCAACCATAGTCCC 180

Qy   181 GCCCCTAACTCCGCCCATCCCGCCCCCAGTTCCCGCCCCA 240
          ||||||||||||||||||||||||||||||||||||||||
Db   181 GCCCCTAACTCCGCCCATCCCGCCCCCAGTTCCCGCCCCA 240

Qy   241 TGGCTGACTAATTTTTTTATTTATGCAGAGGCCGAGGCCTCGAGCTATT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGCTGACTAATTTTTTTATTTATGCAGAGGCCGAGGCCTCGAGCTATT 300

Qy   301 CCAGAAGTACTGAGGAGGCTTTTTTTGCAAAAAGCTTACATGATC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCAGAAGTACTGAGGAGGCTTTTTTTGCAAAAAGCTTACATGATC 360

Qy   361 TGCAGAGAGGCCCAGTATCAGCACTCTCTGCAGTCATGGGCTTCACAGCT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGCAGAGAGGCCCAGTATCAGCACTCTCTGCAGTCATGGGCTTCACAGCT 420

Qy   421 AGCCGGTGACTAGGGCTTAAGATTGGAGCCACCATTAAAGAAGGAAGAAAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCCGGTGACTAGGGCTTAAGATTGGAGCCACCATTAAAGAAGGAAGAAAA 480

Qy   481 AGAAGGAAAGAAAAAAAAAAAAAAAAAGGAAAAAGAAAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||
Db   481 AGAAGGAAAGAAAAAAAAAAAAAAAAAGGAAAAAGAAAAGA 540

Qy   1561 CTTAAAATCGCAGTATCCCGGAATGAATTTGATTCGCCAAAAATAGGAATCTCTGGGCATGCCGAG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CTTAAAATCGCAGTATCCCGGAATGAATTTGATTCGCCAAAAATAGGAATCTCTGGGCATGCCGAG 1620

Qy   1621 AATCTCACGCCAGGCAGTTCTATGAGGCAGAGGCGACACCTTTAGGCAGACCAGTAGATCCA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AATCTCACGCCAGGCAGTTCTATGAGGCAGAGGCGACACCTTTAGGCAGACCAGTAGATCCA 1680

Qy   1681 GAGGAGTTCATGATCAGTCGCAATTGTCTTGTCTTGTTCCCTATCGAAAGAACAAAAATCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GAGGAGTTCATGATCAGTCGCAATTGTCTTGTCTTGTTCCCTATCGAAAGAACAAAAATCG 1740

Qy   1741 TATTCATTAAAAACGGGAGGTAGGATGAGATGAGATGTGACGTGTACATCGACTGAAATCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TATTCATTAAAAACGGGAGGTAGGATGAGATGAGATGTGACGTGTACATCGACTGAAATCCC 1800

Qy   1801 TGGTAATCCGTTTTAGAATCCATGAATAATAATATTTTTTGGGATGGGAAGGCTTTTTTGC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGGTAATCCGTTTTAGAATCCATGAATAATAATATTTTTTGGGATGGGAAGGCTTTTTTGC 1860

Qy   1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTTGGGAAACGAACACCACGGTAGGCTGGGAAATG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ACGTTCAAAATTTTTTGCAACCCCTTTTTTGGGAAACGAACACCACGGTAGGCTGGGAAATG 1920

Qy   1921 CCCATACTGTTTGACCAATTCACGTTCATTTATTATATAATAATGTTAAAATGTCGTTCAACT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CCCATACTGTTTGACCAATTCACGTTCATTTATTATATAATAATGTTAAAATGTCGTTCAACT 1980

Qy   1981 CCGATAAATAAACGGCCCAACACCCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CCGATAAATAAACGGCCCAACACCCGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACG 2040

Qy   2041 ACGATTCTGTGATTTTGTATTCAGCCCATATCGTTTCATRAGCTTCTGCCAACCGAACCGGAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACGATTCTGTGATTTTGTATTCAGCCCATATCGTTTCATRAGCTTCTGCCAACCGAACCGGAC 2100

Qy   2101 ATTTCGAAGTACTCAGCGTAAGTGAATGGTCCACCTCGATAATGTGCATTCTGTAAAAAGCAATT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATTTCGAAGTACTCAGCGTAAGTGAATGGTCCACCTCGATAATGTGCATTCTGTAAAAAGCAATT 2160

Qy   2161 GTTCCAGGAGAACCAGGGCGTATCTCTTCATTGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GTTCCAGGAGAACCAGGGCGTATCTCTTCATTGCCTTATGCAGTTGCTCTCCAGCGGTTCCA 2220

Qy   2221 TCTTCCAGGGGATAGGAATGGCGCCGGGGCCTTTCTTTATGTTTTTTTGGCGTCTTCCATGGGGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TCTTCCAGGGGATAGGAATGGCGCCGGGGCCTTTCTTTATGTTTTTTTGGCGTCTTCCATGGGGA 2280

Qy   2281 CGTCGGTTGGTGTTACGTTGGTTTTCTTTCGGATTCGTGCTCATGATGCAC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CGTCGGTTGGTGTTACGTTGGTTTTCTTTCGGATTCGTGCTCATGATGCAC 2340

Qy   2341 GGTCTTACGAGACCTCCGCGGGGCACTCGCACGAGTACCACACAAAGGCCTT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GGTCTTACGAGACCTCCGCGGGGCACTCGCACGAGTACCACACAAAGGCCTT 2400

Qy   2401 TCGCGGACCCAACACTACTCGGCTAGCAGTCTTGGGGGGGCACGCCCAAATCTCCAGGCAT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TCGCGGACCCAACACTACTCGGCTAGCAGTCTTGGGGGGGCACGCCCAAATCTCCAGGCAT 2460

Qy   2461 TGACGGGGGTTTATCCAAGAAAGGACCCGTCGTTCGTCTTGGCAAATTCGGTGTACTCACCGGT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TGACGGGGGTTTATCCAAGAAAGGACCCGTCGTTCGTCTTGGCAAATTCGGTGTACTCACCGGT 2520

Qy   2521 TCCGGCGAGACCACTATGGCTCTTCCCGGGGAGGGGTTCTCGGAGGGTCGACGACACTCATA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TCCGGCGAGACCACTATGGCTCTTCCCGGGGAGGGGTTCTCGGAGGGTCGACGACACTCATA 2580

Qy   2581 CTAACGCCACCATGGCTAGACCGCTTTTCTGCGTGAAGAGACAGTAGTTCCTCACAGGGAGTGATT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CTAACGCCACCATGGCTAGACCGCTTTTCTGCGTGAAGAGACAGTAGTTCCTCACAGGGAGTGATT 2640

Qy   2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGCGCTGGCCGCCGGCCGGCCTTCCTGC 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGCGCTGGCCGCCGGCCGGCCTTCCTGC 2700

Qy   2701 TGGGCGCCGGCTGGGCAACATTCCGAGGGGGACCGGTCCCGTCGGTGGGGGGG 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TGGGCGCCGGCTGGGCAACATTCCGAGGGGGACCGGTCCCGTCGGTGGGGGGG 2760

Qy   2761 ACAAATCTCTC 2771
          |||||||||||
Db   2761 ACAAATCTCTC 2771

Left column:

```
Qy    541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAAATGGGCTAAGA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAGGAAAATGGCCTAAGA 600

Qy    601 GGCCGGGAGTGTTTACCCCAACCTTTAAACGGCGGGGANCTTTCCGCCCTTTCTTGGGCCTTTATG 660
          ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db    601 GGCCGGGAGTGTTTACCCCAACCTTTAAACGGCGGGATCTTTCCGCCCTTTCTTGGGCCTTTATG 660

Qy    661 AGGATCTCTCTGATTTTTTTGCCTCGAGTTTTTCGGGTTAAGACCTTTGGGTACTTCGTCC 720
          |||||||||   ||||   |||||||||||||||||   ||||
Db    661 AGGATCTCTCTGATTTTTTTGCCTCGAGTTTTTCGGGTTAAGACCTTTGGGTACTTCGTCC 720

Qy    721 ACAAAACACAACTCCTCCCGCGCAACTTTTTTCGGTGTTACTTGACTCGGCGACGTAACTC 780
          |||||||||||||||||||||||||||||||
Db    721 ACAAAACACAACTCCTCCCGCGCAACTTTTTTCGGTGTTACTTGACTCGGCGACGTAACTC 780

Qy    781 ACGATCTCTTTTTCCGTCATCGTCGCGGCAACTTTTTTCGGTGTTTCCGGTGTTTTCGTC 840
          ||||||||||||||||||||||||
Db    781 ACGATCTCTTTTTCCGTCATCGTCGCGGCAACTTTTTTCGGTGTTTCCGGTGTTTTCGTC 840

Qy    841 TCACCGGGGTCATCGTCGGGGAAGCACCTCGCGACACCTGGGTCGAAGATGTTGGGGTGTTGG 900
Db    841 TCACCGGGGTCATCGTCGGGGAAGCACCTCGCGACACCTGGGTCGAAGATGTTGGGGTGTTGG 900

Qy    901 AGCAAGAGATGGGATTCCAATTCAGCGGGGGAGGGCCACCTGGACACCTGGGTCGAAGATTCAGAGAC 960
Db    901 GACATAANCATANGGACCTTCGCACACACAGTTCGCCTCTCGACACAGTTCGCCTCTCTTTGATTAAGCAGGTTTTC 960

Qy    961 TTCAGGCGGGTCAACGATGAAGGAAGTGTTGGGGTTGTCTTCGTCCTCCCAGTAGTTAATCAGAA 1020
Db    961 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCC 1020

Qy   1021 TGTAGCCATCCCATCCTTTCTGTCAATCAAGGCGGTTGGTCGTTCCGGATTGTTTTACATAACG 1080
Db   1021 GGTTTATCATCCCCCTCGGGTGTTAATCAGGCTTGTAGTCTCTCAGTGATCTTTGTAA 1080

Qy   1081 GACATAANCATANGGACCTTCGCACACACAGTTCGCCTCTCGACACAGTTCGCCTCTCTTTGATTAAGCAGGTTTTC 1140
Db   1081 GACATAANCATANGGACCTTCGCACACACAGTTCGCCTCTCGACACAGTTCGCCTCTCTTTGATTAAGCAGGTTTTC 1140

Qy   1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCC 1200
Db   1141 CCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCC 1200

Qy   1201 GGTTTATCATCCCCCTCGGGTGTTAATCAGGCTTGTAGTCTCTCAGTGATCTTTGTAA 1260
Db   1201 CCTTGCCTGATACCTGGGCAGAATGGAACCTCTTGGCAACTTCCTTAGAG 1320

Qy   1261 CCTTGCCTGATACCTGGGCAGAATGGAACCTCTTGGCAACTTCCTTAGAG 1320
Db   1261 CCTTGCCTGATACCTGGGCAGAATGGAACCTCTTGGCAACTTCCTTAGAG 1320

Qy   1321 AGGGGAGGCGCCACCCAGAAGCAATTTCGTGTGTAAAATAGATAAAATCGTATTTGTCAATCAGA 1380
Db   1321 AGGGGAGGCGCCACCCAGAAGCAATTTCGTGTGTAAAATAGATAAAATCGTATTTGTCAATCAGA 1380

Qy   1381 GTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCCAGCGCCACTTTGAATCTTGTAA 1440
Db   1381 GTGCTTTTGGCGAAGAAGGAGAATAGGGTTGGCACCCAGCGCCACTTTGAATCTTGTAA 1440

Qy   1441 TCCTGAAGGCTCCTCAGAAAACAGCTCCTTCTTCTTCAAATCTTATACATTAAGACGACTCGAAAT 1500
Db   1441 TCCTGAAGGCTCCTCAGAAAACAGCTCCTTCTTCTTCAAATCTTATACATTAAGACGACTCGAAAT 1500

Qy   1501 CCACATATCAAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560
Db   1501 CCACATATCAAAATATCCGAGTGTAGTAAACATTCCAAAACCGTGATGGAATGGAACAACA 1560

Qy   1561 CTTAAAATCGCAGTATCCGGAANGAATGATTTGATTGCCAAAAAATAGGATCTCTGGCCATGCCATGGGAG 1620
Db   1561 CTTAAAATCGCAGTATCCGGAANGAATGATTTGATTGCCAAAAAATAGGATCTCTGGCCATGCCATGGGAG 1620
```

Right column:

```
Qy   1621 AATCTCACCCAGGGCAGTTCTTATGAGGGCAGAGGCGACACCTTTAGGCGAGACCAGTAGATCCA 1680
Db   1621 AATCTCACCCAGGGCAGTTCTTATGAGGGCAGAGGCGACACCTTTAGGCGAGACCAGTAGATCCA 1680

Qy   1681 GAGGAGTTCATGAATCAGTGGCAAATNGTCTTCTTGTCCCTATCGGAAGGGACTCTGGCCACAAAATCG 1740
Db   1681 GAGGAGTTCATGAATCAGTGGCAAATNGTCTTCTTGTCCCTATCGGAAGGGACTCTGGCCACAAAATCG 1740

Qy   1741 TATTCATTAAAAACCGGGAGGTAGATGCAAGGTAGAATGTGAACGGTGTACATGGAAATCCC 1800
Db   1741 TATTCATTAAAAACCGGGAGGTAGATGCAAGGTAGAATGTGAACGGTGTACATGGAAATCCC 1800

Qy   1801 TGGGTAAATCCGTTTGTAGAATTCCATGATTAATAATAATTTTTTTGGATGATTGATTGGGGAGACTTTTTTTGC 1860
Db   1801 TGGGTAAATCCGTTTGTAGAATTCCATGATTAATAATAATTTTTTTGGATGATTGATTGGGGAGACTTTTTTTGC 1860

Qy   1861 ACGTTCAAAAATTTTTTGCAACCCCCTTTTTTGGAAACCGAACACCACGGGTAGGCTAGGCGGAAATG 1920
Db   1861 ACGTTCAAAAATTTTTTGCAACCCCCTTTTTTGGAAACCGAACACCACGGGTAGGCTAGGCGGAAATG 1920

Qy   1921 CCCATACTGTTGAGCAATTCACGTTCATTATAAAATGTCGTTCGGTTCGGGGGGGGGGCGAAACTGCAAAT 1980
Db   1921 CCCATACTGTTGAGCAATTCACGTTCATTATAAAATGTCGTTCGGTTCGGGGGGGGGGCGAAACTGCAAAT 1980

Qy   1981 CCGGATAAATAAACGGCGGCCAACACCCGGCATAAAAGAATTGAAAGAGAGTTTTTCAGTCGATACG 2040
Db   1981 CCGGATAAATAAACGGCGGCCAACACCCGGCATAAAAGAATTGAAAGAGAGTTTTTCAGTCGATACG 2040

Qy   2041 ACGATTCGTGATTTGTNGTATTCAGCCCAATATCGTTTCATTCGTTTCCATGAACCGAACGGAC 2100
Db   2041 ACGATTCGTGATTTGTNGTATTCAGCCCAATATCGTTTCATTCGTTTCCATGAACCGAACGGAC 2100

Qy   2101 ATTTCGAAGTACTCAGCGTAAGTGAGTTCACCTCGGATATGTGAATCTGTAAAAGCAAATT 2160
Db   2101 ATTTCGAAGTACTCAGCGTAAGTGAGTTCACCTCGGATATGTGAATCTGTAAAAGCAAATT 2160

Qy   2161 GTTCCAGGAACCCAGGGCGTATCTCTTCATGAGTTTCAGTTGCTCTCCAGGCCGTTCCA 2220
Db   2161 GTTCCAGGAACCCAGGGCGTATCTCTTCATGAGTTTCAGTTGCTCTCCAGGCCGTTCCA 2220

Qy   2221 TCTTCCAGCGGGATAGGAATGGCCTTGGCCTTTCTTAATGTTTTTGGGCGTCGTCTTTAGGGA 2280
Db   2221 TCTTCCAGCGGGATAGGAATGGCCTTGGCCTTTCTTAATGTTTTTGGGCGTCGTCTTTAGGGA 2280

Qy   2281 CGTCGGTTGGTTGTTTACGTTGGTTTTTTTTCTTTGGAGTTTTAGGGATTTCGGTCTGTTACTGCAC 2340
Db   2281 CGTCGGTTGGTTGTTTACGTTGGTTTTTTTTCTTTGGAGTTTTAGGGATTTCGGTCTGTTACTGCAC 2340

Qy   2341 GGTCTACGGAGACCTTCCCGGGGACGACTCGCAAGCACCTTATCGAGGCAAGCAGGGCCTT 2400
Db   2341 GGTCTACGGAGACCTTCCCGGGGACGACTCGCAAGCACCTTATCGAGGCAAGCAGGGCCTT 2400

Qy   2401 TCGGCGGACCCAACACTACTGGGCTAGCAGTCTTGCGGGGGGCCCACGCGCCAAATCTCCAGGCATT 2460
Db   2401 TCGGCGGACCCAACACTACTGGGCTAGCAGTCTTGCGGGGGGCCCACGCGCCAAATCTCCAGGCATT 2460

Qy   2461 TGAGCGGGGGTTATCCAAGAAAGGACACCCGGTCGTGTTACTCACCGGT 2520
Db   2461 TGAGCGGGGGTTATCCAAGAAAGGACACCCGGTCGTGTTACTCACCGGT 2520

Qy   2521 TCCGCAGACCACTATGGCTCTCCCGGGGAGGGGGGGTTCCTGGCGAACACTCATA 2580
Db   2521 TCCGCAGACCACTATGGCTCTCCCGGGGAGGGGGGGTTCCTGGCGAACACTCATA 2580

Qy   2581 CTAACGCCATGGCTAGACGGCCTTTCTGCGTTGAAGACAGTAGTTCCTCACAGGGGAGTGATT 2640
Db   2581 CTAACGCCATGGCTAGACGGCCTTTCTGCGTTGAAGACAGTAGTTCCTCACAGGGGAGTGATT 2640

Qy   2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGGCTGGC 2674
Db   2641 CATGGTGGAGTGTCGCCCCCATCAGGGGGGCTGGC 2674
```

RESULT 4
US-10-066-130-20
; Sequence 20, Application US/10066130
; Patent No. 6699657
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: In Vitro System for Replication of RNA-Dependent RNA Polymerase (
; TITLE OF INVENTION: Viruses
; FILE REFERENCE: PH-7171 NP
; CURRENT APPLICATION NUMBER: US/10/066,130
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/265,437
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 2327
; TYPE: DNA
; ORGANISM: viral
US-10-066-130-20

Query Match   84.0%;   Score 2327;   DB 4;   Length 2327;
Best Local Similarity   100.0%;
Matches 2327;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    348 AGCTTACATGATCTGCAGAGAGGAGGCCAGTAATCAGACACTCTCTGCAGTCATGCGGCTCACGG 407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGCTTACATGATCTGCAGAGAGGAGGCCAGTAATCAGACACTCTCTGCAGTCATGCGGCTCACGG 60

Qy    408 ACCTTTCACAGCTAGCCGTGACTAGGGCTAAGATGGAGCCACCAATTAAAGAAAGGAAGGAA 467
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACCTTTCACAGCTAGCCGTGACTAGGGCTAAGATGGAGCCACCAATTAAAGAAAGGAAGGAA 120

Qy    468 AAGAAAGGAAGAAAAAGAGAAAGAGAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA 527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAGAAAGGAAGAAAAAGAGAAAGAGAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180

Qy    528 AAGAAAAAAAAAAAAAAAA 587
          |||||||||||||||||||
Db    181 AAGAAAAAAAAAAAAAAAA 240

Qy    588 AAAATGGCCCTAAGAGCCCGGAGTGTTTAACCCAACCTTTTCCGCCCCTT 647
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAAATGGCCCTAAGAGCCCGGAGTGTTTAACCCAACCTTTTCCGCCCCTT 300

Qy    648 CTTGGCCTTTATGAGGATCTCTCTGATTTTCTTGCGTTGAGTTTTCCGGTAAGAACCTTT 707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTTGGCCTTTATGAGGATCTCTCTGATTTTCTTGCGTTGAGTTTTCCGGTAAGAACCTTT 360

Qy    708 CGGTACTTCGTCCACAAACACAACTCCCCGCGCAAACTTTTTGTTTACTTGACT 767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGGTACTTCGTCCACAAACACAACTCCCCGCGCAAACTTTTTGTTTACTTGACT 420

Qy    768 GGCGACGTAATCCAGATCTCTTTTTCCGTCTCTTTCCGTCCTCCAACAACAACAAC 827
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGCGACGTAATCCAGATCTCTTTTTCCGTCTCTTTCCGTCCTCCAACAACAACAAC 480

Qy    828 GGCCGCGGCGGGAGAAGTTCACCGCGCGTCATCGTGCGTGCGTTGCGGAAGAT 887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGCCGCGGCGGGAGAAGTTCACCGCGCGTCATCGTGCGTGCGTTGCGGAAGAT 540

Qy    888 GTTCGGGGTGTTGGGAGGACAAGATGGGATTCCAATTCAGCGGGGAGCCACCTGATTTGTA 947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTCGGGGTGTTGGGAGGACAAGATGGGATTCCAATTCAGCGGGGAGCCACCTGATTTGTA 600

Qy    948 CTTAATCAGAGACTTCAGGCGCGGTCAACGATGGGAAGAGTGTTCGTCTTTCGTTCCCAGTAAGC 1007
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAATCAGAGACTTCAGGCGCGGTCAACGATGGGAAGAGTGTTCGTCTTTCGTTCCCAGTAAGC 660

Qy   1008 TATGTCTCCAGAAAGTAGCCATCCATCCTTTGTCAATCAAGGCGTTGGGTCGCTTCCGGATT 1067
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TATGTCTCCAGAAATGTAGCCATCCATCCTTTGTCAATCAAGGCGTTGGGTCGCTTCCGGATT 720

Qy   1068 GTTTACATAACCGGACTAATCATAGGACCTTCTCACACACAGTTCGCCCTCTTTGATTAAC 1127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTTTACATAACCGGACTAATCATAGGACCTTCTCACACACAGTTCGCCCTCTTTGATTAAC 780

Qy   1128 GCCCAGCCGTTTTCCCGGGTATCCACAGATCCACAACCTTCGCTTCAAAAAATGGAACACTTT 1187
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCCAGCCGTTTTCCCGGGTATCCACAGATCCACAACCTTCGCTTCAAAAAATGGAACACTTT 840

Qy   1188 ACCGGACCGCGCGCCCGGGTTTATCATCCCCCTCGGGGTTGGGTGTAATCAGAAGCTGATGTGTCTC 1247
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACCGGACCGCGCGCCCGGGTTTATCATCCCCCTCGGGGTTGGGTGTAATCAGAAGCTGATGTGTCTC 900

Qy   1248 AGTGAGCCCATACCTGCCTGATACCTGGCAGATGGAACCTTCCCC 1307
          |||||||||||||||||||||||||||||||||||||||||||||
Db    901 AGTGAGCCCATACCTGCCTGATACCTGGCAGATGGAACCTTCCCC 960

Qy   1308 GACTTCCTTAGAGGAGGGGGACGCGCCACCCAGAAGCAAATTTCGTGTAAAATTAGATAAATCGTA 1367
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GACTTCCTTAGAGAGGGGACGCGCCACCCAGAAGCAAATTTCGTGTAAAATTAGATAAATCGTA 1020

Qy   1368 TTTGTCAATCAGAGTGCTTTTGGCGAAGAGGAGGAGAGAATAGGGTTGGGCACCAGCAGCCACT 1427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TTTGTCAATCAGAGTGCTTTTGGCGAAGAGGAGGAGAGAATAGGGTTGGCACCAGCAGCCACT 1080

Qy   1428 TTGAATCTTGTAATCCTGAAGGCTCCTCAGAAAACAGCTCTTCTTTCTTCAAAATCTTATACATTTAA 1487
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGAATCTTGTAATCCTGAAGGCTCCTCAGAAAACAGCTCTTCTTTCTTCAAAATCTTATACATTTAA 1140

Qy   1488 GACGACTCGAAATCCACATATCAAATATCGGAGTGTTAGTAAACAATTCCAAAACCCGTGATG 1547
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACGACTCGAAATCCACATATCAAATATCGGAGTGTTAGTAAACAATTCCAAAACCCGTGATG 1200

Qy   1548 GAATGGAACAACACTTAAAATCGGAAATCGCAGTATCCGAGTATCCGAGGAATGGAATCGGAATC 1607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAATGGAACAACACTTAAAATCGGAAATCGCAGTATCCGAGTATCCGAGGAATGGAATCGGAATC 1260

Qy   1608 TCTGCCATGGGAGAATCTCACGCAGGCAGTTCTATGAGGCAGACCTTTAGGCAG 1667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TCTGCCATGGGAGAATCTCACGCAGGCAGTTCTATGAGGCAGACCTTTAGGCAG 1320

Qy   1668 ACCAGTAGATCCAGAGAGGAGTTCATGAGAGTCATGATCAGTGAGTGCAGGGACTC 1727
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ACCAGTAGATCCAGAGAGGAGTTCATGAGAGTCATGATCAGTGAGTGCAGGGACTC 1380

Qy   1728 TGGCACAAAATCGTATTCATTGAAATCGTGGTTCATTAATAAAATGTCGTTGCGGG 1787
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGCACAAAATCGTATTCATTGAAATCGTGGTTCATTAATAAAATGTCGTTGCGGG 1440

Qy   1788 CGACTGAAATCCCTGGGTAATCCGTTTTAGAATCCATGATAAATAATTTTTTGGGATGATTGG 1847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CGACTGAAATCCCTGGGTAATCCGTTTTAGAATCCATGATAAATAATTTTTTGGGATGATTGG 1500

Qy   1848 GAGCTTTTTTGCAGTTCAAAAATTTTTGCAACCCCTTTTTGGAAAACGAACACCACGGT 1907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAGCTTTTTTGCAGTTCAAAAATTTTTGCAACCCCTTTTTGGAAAACGAACACCACGGT 1560

Qy   1908 AGGCTGCGAAATGCCCATACTGTTGAGCAATTCAGTTCATTCATTATAAATGTCGTTGCGGG 1967
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGGCTGCGAAATGCCCATACTGTTGAGCAATTCAGTTCATTCATTATAAATGTCGTTGCGGG 1620

Qy   1968 CGCACTGCAACTCCGATAAAATAACGCGCCCAACACCGGCATAAAGAAAGAAGTTGAAGAGAGTT 2027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CGCACTGCAACTCCGATAAAATAACGCGCCCAACACCGGCATAAAGAAAGAAGTTGAAGAGAGTT 1680

Qy   2028 TTCACTCGATACGACGGATTCGTGTGATTTGTTATTCAGCCCATATCGTTCAAGTACTTCTGC 2087
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TTCACTCGCATACGACGGATTCTGTGATTTGTTATTCAGCCCATATCGTTCAAGTACTTCTGC 1740

Qy   2088 CAACCGAACGGACACTTTCGAAGTACTCACGGTAAGTGATGTCCACCTCGATTATGTTGCATC 2147
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CAACCGAACGGACACTTTCGAAGTACTCACGGTAAGTGATGTCCACCTCGATTATGTTGCATC 1800

Qy   2148 TGTAAAAAGCAATTGTTCCAGGAACCAGGGGGTATCTCTTCAGATAGCCTTATGCAGTTGCTC 2207
Db
```

RESULT 5
US-09-743-194-28/c
Sequence 28, Application US/09743194
Patent No. 6716601
GENERAL INFORMATION:
APPLICANT: Belfield, Graham
APPLICANT: Oakley, Caroline
TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast
FILE REFERENCE: 3526.82643
CURRENT APPLICATION NUMBER: US/09/743,194
CURRENT FILING DATE: 2001-01-08
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 13654
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-743-194-28

```
Query Match       59.8%;  Score 1656.4;  DB 4;  Length 13654;
Best Local Similarity  98.8%;  Pred. No. 2.5e-303;
Matches 1669;  Conservative  0;  Mismatches 21;  Indels  0;  Gaps  0;
```

(Sequence alignment — Qy = query, Db = database subject)

Left column:

```
Db  1801  TGTAAAAGCAAATTGTTCCAGGAGGACCAGGCGGTACTCTTCCATTATGCAGTTGCTC  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2208  TCCAGCGGTTCCATCTTCCAGGGGATAGAAATGGGCGCCGGGCCTTTCTTCTTTTGGC  2267
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  TCCAGCGGTTCCATCTTCCAGCGGGATAGAAATGGGCGCCGGGCCTTTCTTTTTTGGC  1920
Qy  2268  GTCTTCCATGGGGACGTCGGTTGGTGTGTTAGGTTGTTAGGTTAGGATTCGT  2327
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  GTCTTCCATGGGGAGGTCGGTTGGTTGGTTTTCTTTGAGGTTAGGATTCGT  1980
Qy  2328  GCTCATGATGCACGGTCTTACGGAGCTCCGGAGGCACCCTATCAGGCAGT  2387
Db  1981  GCTCATGATGCACGGTCTTACGGAGCTCCGGAGGCACCCTATCAGGCAGT  2040
Qy  2388  ACCACAAGGCCTTTCGCGACCACCACTACTCGCTGGGGGGCACGCCCA  2447
Db  2041  ACCACAAGGCCTTTCGCGACCCAACACTACTCGCTGGGGGGCACGCCCA  2100
Qy  2448  AATCTCCAGGCCATTGAGCGGGGGTTTATCCAAGAGAAAGGACCCGGGTCGTCGTCC  2507
Db  2101  AATCTCCAGGCATTGAGCGGGGGTTATCCAAGAGAAAGGACCCGGGTCGTCGTCC  2160
Qy  2508  TGTACTCACCGGTTCCGCAGAACCACTATGGCTCTCCCGGGAGGGGGG  2567
Db  2161  TGTACTCACCGGTTCCGCAGAACCACTATGGCTCTCCCGGGAGGGGGG  2220
Qy  2568  CACGACACTCATACTACTAAGCGCCATGGCTTAGACGCTTTCTGCGTGA  2627
Db  2221  CACGACACTCATACTAAGCGCCATGCCTAGACGCTTTCTGCGTGA  2280
Qy  2628  CAGGGGAGTGATTCATGGTGGGAGTGTCGCCCCCATCAGGGGGCTGGC  2674
Db  2281  CAGGGGAGTGATTCATGGTGGGAGTGTCGCCCCCATCAGGGGGCTGGC  2327
```

Right column:

```
Qy  2688  CTTTCCGTGCTCCAAAAACAACAACGGGCGCGGGAAGTTCACCGGGCGGGAAG  2629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   864  ACCTGCGACACCTGCGTCGAAGATGTTGGGGGTGTTGGAGGCAAGGATTCCAATTCAGC  923
Db  2628  ACCTGCGACACCTGCGGTCGAAGATGTTGGGGGTGTTGGAGGCAAGGATTCCAATTCAGC  2569
Qy   924  GGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGTCAACGATGAAGAA  983
Db  2568  GGGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGTCAACGATGAAGAA  2509
Qy  1464  CTCTTCTTCAAATCTATCACATTAAGACGACTCGAAATCCACTATCAAATATCCGAGTGT  1523
Db  2028  CTCTTCTTCAAATCTATCACATTAAGACGACTCGAAATCCACTATCAAATATCCGAGTGT  1969
Qy  1404  TAGGGTTGGCACCACCAGCCCACTTTGAATCTTGTAATCCTGAGGGTCTTCAGAAACAG  1463
Db  2088  TAGGGTTGGCACCACCAGCCCACTTTGAATCTTGTAATCCTGAGGGTCTTCAGAAACAG  2029
Qy  1344  TTCGTGTAAATTAGGATAAATCGTATTTTGTCAATCAGAGTGCTTTTGGCGAAGAAGGAGA  1403
Db  2148  TTCGTGTAAATTAGGATAAATCGTATTTTGTCAATCAGAGTGCTTTTGGCGAAGAAGGAGA  2089
Qy  1284  GAACCTTCTTGGGCAACCGCTTCCCGGACTTCCTTAGAGAGGGGGGCCGCCAGAAGCAAT  1343
Db  2208  GAACCTTCTTGGGCAACCGCTTCCCGGACTTCCTTAGAGAGGGGGGCCGCCAGAAGCAAT  2149
Qy  1164  CGCTTCAAAAATGGAACAACTTTACCGACCGGCGGCCGGGTTATCATCCCCTCGGGTGT  1223
Db  2388  CGCTTCAAAAATGGAACAACTTTACCGACCGGCGGCCGGGTTATCATCCCCTCGGGTGT  2329
Qy  1104  ACACAGTTCGCCTCTTTGATTAAACGCCCAGCCGTTTTCCCGGTATCCAGATCCACAACTT  1163
Db  2388  ACACAGTTCGCCTCTTTGATTAAACGCCCAGCCGTTTTCCCGGTATCCAGATCCACAACTT  2329
Qy  1044  CAAGGCGTTGGTCGTTCCGGATTGTTTACATAACCGGACATAATCATAGGACCTCTCAC  1103
Db  2448  CAAGGCGTTGGTCGTTCCGGATTGTTTACATAACCGGACATAATCATAGGACCTCTCAC  2389
Qy  1644  AGGCAGGAGGCGACACCTTAGGCAGACCAGTAGATCCAGGAGGAGTTCATGATCAGTGCAAT  1703
Db  1848  AGGCAGGAGGCGACACCTTAGGCAGACCAGTAGATCCAGGAGGAGTTCATGATCAGTGCAAT  1789
Qy   744  CTTTTTCGCGGGTTCGTTACTTGACTCGGCGACGTAATCCAGATCTCTTTTCCGTCATCGT  803
Db  2748  CTTTTTCGGCGGGTTTCGGGTTTTGGGATGATTGGGAGCTTTTTTGCACGTTCAAACCC  2689
Qy   804  CTTTCCGTGCTCCAAAAACAACAACGGGCGCGGGAAGTTCACCGGGCGGGAAG  863
```

Db 1608 CTTTTTGGAAACAGACACCACGGTAGGCTCGGCGAAATGCCATACTGTTGAGCAATTCACG 1549
Qy 1944 TTCATTATAAAATGTGTTCCGGGGGGCAACTTGCACACTCGGATAAATAATAACCGCCAACAC 2003
Db 1548 TTCATTATAAAATGTGTTCCGGGGGGCAACTTGCACACTCGGATAAATAATAACCGCCAACAC 1489
Qy 2004 CGGCATAAAAGAAATTGAAGAGAGAGTTTTCACTGCCGTATCGTACGACGATTCGTGATTTGATTTGTATTCAG 2063
Db 1488 CGGCATAAAAGAAATTGAAGAGAGAGTTTTCACTGCCGTATCGTACGACGATTCGTGATTTGATTTGTATTCAG 1429
Qy 2064 CCCATATCGTTTCATAGCCTTCTGCCAACCGAACGGACACATTTGCCGTATCGTACTACTCAGCGTAAGT 2123
Db 1428 CCCATATCGTTTCATAGCCTTCTGCCAACCGAACGGACACATTTGCCGTATCGTACTACTCAGCGTAAGT 1369
Qy 2124 GATGTCCACCTCGATATGTGCCATCGTGTATAAAAGCAATTGTTCCAGGAACCAGGGCGTATCT 2183
Db 2184 CTTCATAGCCTTATGCAGTTGCTCTTCCACGCGGTTCCATCTTCTCCAGCGGGTTCCATCTTTGATTAACGCC 2243
Qy 2244 CGGGCCTTTCTTTATGTTTTTGGCGTCTTCCAGACTGTCGGGTTGGTTTACGTTGGT 2303
Db 1248 CGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATATGTGCTGTTGTTAGAATTAGGTTTAGCT 1189
Qy 2304 TTTTCTTTGA 2313
Db 1188 TGTTTTTTTA 1179

RESULT 6
US-09-602-628-7/c
; Sequence 7, Application US/09602628
; Patent No. 6495355
; GENERAL INFORMATION:
; APPLICANT: Eames, Brian
; APPLICANT: Contag, Christopher
; TITLE OF INVENTION: Red-Shifted Luciferase
; FILE REFERENCE: SUN-127
; CURRENT APPLICATION NUMBER: US/09/602,628
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/140,598
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: pGL3
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1650)
US-09-602-628-7

Query Match   59.7%; Score 1655.6; DB 4; Length 1686;
Best Local Similarity   98.9%; Pred. No. 2.3e-303;
Matches 1667; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Db 1506 GACGTAATCCACGATCTCTTTTTCCGTCATCGTCTTTTCCGTGCTCCAAAACAACAAGGC 1447
Qy 831 GGCGGGAAGTTCACCGGCGTCATCGTCGGGAAGACCTGCGTCGAAGAGATGTT 890
Db 1446 GGCGGGAAGTTCACCGGCGTCATCGTCGGGAAGACCTGCGTCGAAGAGATGTT 1387
Qy 891 GGGGTCGTTGGAGCAGAAGATGGATTCCAAATTCAGCGGGAGCGCCAACCTGATAGCCTTTGTACTT 950
Db 1386 GGGGTCGTTGGAGCAGAAGATGGATTCCAAATTCAGCGGGAGCGCCAACCTGATAGCCTTTGTACTT 1327
Qy 951 AATCAGGACTTCAGGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTCTTCCAGTAAGCTAT 1010
Db 1326 AATCAGGACTTCAGGCGGGTCAACGATGAAGAAGTGTTCGTCTTCGTCTTCCAGTAAGCTAT 1267
Qy 1011 GTCTCCAGAATGTAGCCATCCATCCTCTTGTCAATCAAGCCGTTCCCGGGGATTGTT 1070
Db 1266 GTCTCCAGAATGTAGCCATCCATCCTCTTGTCAATCAAGCCGTTCCCGGGGATTGTT 1207
Qy 1071 TACATAACCGGACATAATCATAGGACCTTCTCACACAGCAGGCTTCGCCTTCTTGATTAACGCC 1130
Db 1206 TACATAACCGGACATAATCATAGGACCTTCTCACACAGCAGTTCGCCTTCTTGATTAACGCC 1147
Qy 1131 CAGCGTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACC 1190
Db 1146 CAGCGTTTCCCGGTATCCAGATCCACAACCTTCGCTTCAAAAAATGGAACAACTTTACC 1087
Qy 1191 GACCCGGCGCCCGGTTTATCATCCCCTCGGGTGTAATCAGAATAGTCAGTGTTGACTCCAGT 1250
Db 1086 GACCCGGCGCCCGGTTTATCATCCCCTCGGGTGTAATCAGAATAGTCAGTGTTGACTCCAGT 1027
Qy 1251 GAGCCGGTAATCCTTCGCCTGATACCTGGGCACGAGAATGGAACCTCTTGGCCAAGTCCCGAC 1310
Db 1026 GAGCCGGTAATCCTTCGCCTGATACCTGGGCACGAGAATGGAACCTCTTGGCCAAGTCCCGAC 967
Qy 1311 TTCCCTTAGAGAGGGGAGCGCGCCACCAGAAGCAATTCGTGTAAATTAGATAAATCGTATTT 1370
Db 966 TTCCCTTAGAGAGGGGAGCGCGCCACCAGAAGCAATTCGTGTAAATTAGATAAATCGTATTT 907
Qy 1371 GTCAATCAGAGTGCTTTTGGCGAAGAAGGAGAAATAGGGTGGCACCGGGTCACGAAGACACTTTG 1430
Db 906 GTCAATCAGAGTGCTTTTGGCGAAGAAGGAGAAATAGGGTGGCACCGGGTCACGAAGACACTTTG 847
Qy 1431 AATCTTGTAATCCTGAAGGGCTCCTCAGAAAACAGCTCTTCTTCACAAATCTATACATTAAGAC 1490
Db 846 AATCTTGTAATCCTGAAGGGCTCCTCAGAAAACAGCTCTTCTTCACAAATCTATACATTAAGAC 787
Qy 1491 GACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAAACACGTGATGGAA 1550
Db 786 GACTCGAAATCCACATATCAAATATCCGAGTGTAGTAAAACACGTGATGGAA 727
Qy 1551 TGGACAAACACTTAAAATCCGAGTATCCGGAATGAATCCGAGTGTAAAAAATGATCTCT 1610
Db 726 TGGACAACAACACTTAAAATCCGAGTATCCGGAATGAATCCGAGTGTAAAAAATGATCTCT 667
Qy 1611 GGCATGCGAGGAAATCCACGGCAGGCAGTTCTATGAGGCAGCGGACGACCTTAGGCAGACC 1670
Db 666 GGCATGCGAGGAAATCCACGGCAGGCAGTTCTATGAGGCAGCGGACGACCTTAGGCAGACC 607
Qy 1671 AGTAGATCCAGAGGAGTTCAATGATCAGTGCAAATTGTCTTGTCCCTATCGAAGGAGTCTGG 1730
Db 606 AGTAGATCCAGAGGAGTTCAATGATCAGTGCAAATTGTCTTGTCCCTATCGAAGGAGTCTGG 547
Qy 1731 CACAAAATCTGTATTCATTAAAACCGGGAGTGAGATGTGACGAACGTGTACATCGA 1790
Db 546 CACAAAATCTGTATTCATTAAAACCGGGAGTGAGATGTGACGAACGTGTACATCGA 487
Qy 1791 CTGAAATCCCTGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGAGTGATGGGGAG 1850
Db 486 CTGAAATCCCTGGTAATCCGTTTTAGAATCCATGATAATAATTTTTTGGAGTGATGGGGAG 427
Qy 1851 CTTTTTTTGCACGTTCACAGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACGAACACCACGGTAGG 1910
Db 426 CTTTTTTTGCACGTTCACAGTTCAAAATTTTTTGCAACCCCTTTTTGGAAACGAACGAACACCACGGTAGG 367
Qy 771 GACGTAATCCACGATCTCTTTTTCCGTCATCGTCTTTTCCGTGCTCCAAAACAACAAGGC 830
Db 1566 TACTTCGTCCACAAAACAACAACTCCCGCGGGCAACTTTTTCGGCGTTGTTACTTGACTGGC 1507

RESULT 7
US-09-743-194-22/c
Sequence 22, Application US/09743194
Patent No. 6716601
GENERAL INFORMATION:
APPLICANT: Belfield, Graham
APPLICANT: Oakley, Caroline
TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast
FILE REFERENCE: 3526.82543
CURRENT APPLICATION NUMBER: US/09/743,194
CURRENT FILING DATE: 2001-01-08
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 22
LENGTH: 12850
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-743-194-22

Query Match   59.7%; Score 1655; DB 4; Length 12850;
Best Local Similarity   98.5%; Pred. No. 4.6e-303;
Matches 1670; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

[Sequence alignment data — Qy/Db rows omitted as dense nucleotide sequence]

```
Qy    2004 CGGGCATAAAAGAATTGAAGAGAGAGAGTTTTCACTCATTACGACGATTCTGTGATTGTGATTGTTATTCAG 2063
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     788 CGGGCATAAAAGAATTGAAGAGAGAGAGTTTTCACTCATTACGACGATTCTGTGATTGTGATTGTTATTCAG  729

Qy    2064 CCCATATCGTTTCATTAGCTTCTGCCAACCGAACCGGACATTTCGAAGTACTTCAGCGGCTAAGT 2123
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     728 CCCATATCGTTTCATTAGCTTCTGCCAACCGAACCGGACATTTCGAAGTACTTCAGCGGCTAAGT  669

Qy    2124 GATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGGGTATCT 2183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     668 GATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGGGTATCT  609

Qy    2244 CGGGGCCTTTCTTTATGTTTTTGGGCGTTCTTCCATGGGGACGTCGGGTTGGGTTTACGTTTTGGT 2303
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     548 CGGGGCCTTTCTTTATGTTTTTGGGCGTTCTTCCATGGGGACGTCGGGTTGGGTTTACGTTTTGGT  489

Qy    2304 TTTTCTTTGAGGTTT 2318
           |||||||||||||||
Db     488 ATGTTTCTTTGATTT  474
```

RESULT 8
US-09-577-424-3/c
; Sequence 3, Application US/0957424
; Patent No. 6525245
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Rhoads, David M
; TITLE OF INVENTION: METHOD FOR IDENTIFYING COMPONENTS INVOLVED IN SIGNAL
; FILE REFERENCE: UNL2990
; CURRENT APPLICATION NUMBER: US/09/577,424
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 14194
; TYPE: DNA
; ORGANISM: Bacterial Plasmid DNA-Plasmid pZB-ALN
US-09-577-424-3

Query Match   59.7%; Score 1654.2; DB 4; Length 14194;
Best Local Similarity   99.8%; Pred. No. 6.7e-303;
Matches 1656; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy     624 TTAAACGGCCATCTTTCCGCGCCCTTCTTATGAGGATCTCTCGATTTTTCTTGC  683
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13826 TTACACGGCGGATCTTTCCGCGCCCTTCTTATGAGGATCTCTCGATTTTTCTTGC 13767

Qy     684 GTCCGAGTTTTCCGGTAAGACCTTTCGGGTACTTCGTCCACAAACACAACTCCTCCGCGCAA  743
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13766 GTCCGAGTTTTCCGGTAAGACCTTTCGGGTACTTCGTCCACAAACACAACTCCTCCGCGCAA 13707

Qy     744 CTTTTTTCGCGGTGTTGTTACTTGACTGACCACGATCTCTTTTTCCGTGTCATCGT  803
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13706 CTTTTTTCGCGGTGTTGTTACTTGACTGGCCGACGTAATCCACGATCTCTTTTTCCGTGTCATCGT 13647

Qy     804 CTTTCCGTGTCTCCAAAAACAACGGCCGCGGGGGGGGGGAGTTCACCGTCGGGGAG  863
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13646 CTTTCCGTGTCTCCAAAAACAACGGCCGCGGGGGGGGGGAGTTCACCGTCGGGGAG 13587

Qy     864 ACCTGCGACACCTGCGCTCGGAGAGATGTTGGGGGGTTGGAGCAAGATGGATTCCAATTCAGC  923
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13586 ACCTGCGACACCTGCGCTCGGAGAGATGTTGGGGGGGTTGGAGCAAGATGGATTCCAATTCAGC 13527

Qy     924 GGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCCGGCTCAACGATGAAGAA  983
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13526 GGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCCGGCTCAACGATGAAGAA 13467

Qy     984 GTGTTGCGTTCTTCGTCCCAGTAAGCTATGTGTCTCCAGAAATGTAGCCATCCATCCTTGTCAAT 1043
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13466 GTGTTCGTCTTCGTCGTCCCAGTAAGCTATGTGTCTCCAGAAATGTAGCCATCCATCCTTGTCAAT 13407

Qy    1044 CAAGGGGTTGGTCGCTTCCCGGATTGTGTTTACATAAACCGGACATAATCATAGGACCTTCTCAC 1103
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13406 CAAGGGGTTGGTCGCTTCCCGGATTGTGTTTACATAAACCGGACATAATCATAGGACCTTCTCAC 13347

Qy    1104 ACACAGTTCGCCCTCTTTGATTAACGCCCCAGCGGTTTTCCCGGTTATCCGACGATCACAACCTT 1163
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13346 ACACAGTTCGCCCTCTTTGATTAACGCCCCAGCGGTTTTCCCGGTTATCCGACGATCACAACCTT 13287

Qy    1164 CGCTTCAAAAAAATGGAACAACTTTACCGACCGGCGCGGCGTTTTCCCGGTTATCCTCGGGTGT 1223
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13286 CGCTTCAAAAAAATGGAACAACTTTACCGACCGGCGCGGCGTTTTCCCGGTTATCCTCGGGTGT 13227

Qy    1224 AATCAGAATAGCTGATGTAGTCTCAGTGAGCCCCATATCCTTGCCTGATACCTGGCAGATG 1283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13226 AATCAGAATAGCTGATGTAGTCTCAGTGAGCCCCATATCCTTGCCTGATACCTGGCAGATG 13167

Qy    1284 GAAACCTCTTGGCCAACCGGCTTCCCCGACTTCCTTAGAGAGGAGGGAGCGCCCACCAGAAGCAAT 1343
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13166 GAAACCTCTTGGCCAACCGGCTTCCCCGACTTCCTTAGAGAGGAGGGAGCGCCCACCAGAAGCAAT 13107

Qy    1344 TTCGTGTAAATTAGATAAAATCGTATTTGTCAAATCAGAGTGGCTTTTGGGCGGAAGAAGGAGAA 1403
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13106 TTCGTGTAAATTAGATAAAATCGTATTTGTCAAATCAGAGTGGCTTTTGGGCGGAAGAAGGAGAA 13047

Qy    1404 TAGGGTTGGCACCAGCAGCCACTTTGAATCTTGTAATCCGAGGCTCCTCAGAAACAG 1463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13046 TAGGGTTGGCACCAGCAGCCACTTTGAATCTTGTAATCCGAGGCTCCTCAGAAACAG 12987

Qy    1464 CTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATATCCGAGTGT 1523
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12986 CTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATATCCGAGTGT 12927

Qy    1524 AGTAAACATTCCAAAACCGTGATGAAATGAAATGAACACACTTTAAAATCGCAGTTCTATG 1583
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12926 AGTAAACATTCCAAAACCGTGATGAAATGAACAACTTTAAAATCGCAGTTATCCGGAAT 12867

Qy    1584 GATTTGATTGCCAAAAAATAGGATCTCTCTGGCATGCAGCGGCAGTTCTATG 1643
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12866 GATTTGATTGCCAAAAAATAGGATCTCTCTGGCATGCAGCGGCAGTTCTATG 12807

Qy    1644 AGGGCAGGAGCGACGACCTTAGGCGAGACCAGTAGATCGCGAGGGAGTTCATGATCAGTGCAAT 1703
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12806 AGGGCAGGAGCGACGACCTTAGGCGAGACCAGTAGATCGCGAGGGAGTTCATGATCAGTGCAAT 12747

Qy    1704 TGTCGTTGTCCCTATCGAAGGAGACTCTGGCACAAAAATCGTATTCATTGAAATCGTAT 1763
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12746 TGTCGTTGTCCCTATCGAAGGAGACTCTGGCACAAAAATCGTATTCATTGAAATCGTAG 12687

Qy    1764 ATGAGAGTGACGAACGTGTACATGACTGAAATCCCTGGTAATCCGTTTGTAGAATCCAT 1823
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12686 ATGAGAGTGACGAACGTGTACATGACTGAAATCCCTGGTAATCCGTTTGTAGAATCCAT 12627

Qy    1824 GATAATAATTTTTTGGATGATGGGAGCTTGTTGCACGTTCAAAAATTTTTTTGCAAACCC 1883
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12626 GATAATAATTTTTTGGATGATGGGAGCTTGTTGCACGTTCAAAAATTTTTTTGCAAACCC 12567

Qy    1884 CTTTTTGGAAAACGAACACACACGGTAGGCTCGGAAATGCCCATACTGTTGAGCAATTCACG 1943
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12566 CTTTTTGGAAAACGAACACACACGGTAGGCTCGGAAATGCCCATACTGTTGAGCAATTCACG 12507

Qy    1944 TTCATTATATAAATGTCGTTCGCGGGCGGGCAACTGCAACTCGGATAAATAACGCCCAAACAC 2003
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12506 TTCATTATATAAATGTCGTTCGCGGGCGGGCAACTGCAACTCGGATAAATAACGCCCAAACAC 12447

Qy    2004 CGGCATAAAAGAATTGAAGAGAGAGAGTTTTCACTGCATACGACGATTCTGTGATTGTATTCAG 2063
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12446 CGGCATAAAAGAATTGAAGAGAGAGTTTTCACTGCATACGACGATTCTGTGATTGTATTCAG 12387

Qy    2064 CCCATATCGTTTCATTAGCTTCTGCCAACCGAACCGGACATTTCGAAGTACTTCAGCGGCTAAGT 2123
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2064
```

```
Db      12386 CCCATATCGTTTCATAGCTTCTGCCAACCGAACGGACATTTCGAAGTACTCAGCGTAAGT 12327

Qy       2124 GATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGCGTATCT 2183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12326 GATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGCGTATCT 12267

Qy       2184 CTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCATCTTCCAGCGGATAGAATGGCGC 2243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12266 CTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCATCTTCCAGCGGATAGAATGGCGC 12207

Qy       2244 CGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGGGACG 2282
              ||||||||||||||||||||||||||||||||||| | |
Db      12206 CGGGCCTTTCTTTATGTTTTTGGCGTCTTCCATGGAAAG 12168

RESULT 9
US-09-577-424-1/c
; Sequence 1, Application US/09577424
; Patent No. 6525245
; GENERAL INFORMATION:
;  APPLICANT:  Rhoads, David M
;  TITLE OF INVENTION:  METHOD FOR IDENTIFYING COMPONENTS INVOLVED IN SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAYS IN HIGHER PLANTS
;  FILE REFERENCE:  UNL2990
;  CURRENT APPLICATION NUMBER: US/09/577,424
;  CURRENT FILING DATE:  2000-05-22
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 12614
;   TYPE: DNA
;   ORGANISM: Bacterial Plasmid DNA - Plasmid pZP-ALE
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1665)..(3317)
US-09-577-424-1

  Query Match             59.7%;  Score 1653.4;  DB 4;  Length 12614;
  Best Local Similarity   99.9%;  Pred. No. 9.2e-303;
  Matches 1654;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        624 TTAAACGGCGATCTTTCCGCCCTTCTTGGCCTTTATGAGGATCTCTCTGATTTTTCTTGC 683
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3317 TTACACGGCGATCTTTCCGCCCTTCTTGGCCTTTATGAGGATCTCTCTGATTTTTCTTGC 3258

Qy        684 GTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCCACAAACACAACTCCTCCGCGCAA 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3257 GTCGAGTTTTCCGGTAAGACCTTTCGGTACTTCGTCCACAAACACAACTCCTCCGCGCAA 3198

Qy        744 CTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCCACGATCTCTTTTTCCGTCATCGT 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3197 CTTTTTCGCGGTTGTTACTTGACTGGCGACGTAATCCACGATCTCTTTTTCCGTCATCGT 3138

Qy        804 CTTTCCGTGCTCCAAAACAACAACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGGAAG 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3137 CTTTCCGTGCTCCAAAACAACAACGGCGGCGGGAAGTTCACCGGCGTCATCGTCGGGAAG 3078

Qy        864 ACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGGAGCAAGATGGATTCCAATTCAGC 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3077 ACCTGCGACACCTGCGTCGAAGATGTTGGGGTGTTGGAGCAAGATGGATTCCAATTCAGC 3018

Qy        924 GGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGTCAACGATGAAGAA 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3017 GGGAGCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGTCAACGATGAAGAA 2958

Qy        984 GTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAAT 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2957 GTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAAT 2898

Qy       1044 CAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCGGACATAATCATAGGACCTCTCAC 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2897 CAAGGCGTTGGTCGCTTCCGGATTGTTTACATAACCGGACATAATCATAGGACCTCTCAC 2838
```

```
Qy       1104 ACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTCCCGGTATCCAGATCCACAACCTT 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2837 ACACAGTTCGCCTCTTTGATTAACGCCCAGCGTTTTCCCGGTATCCAGATCCACAACCTT 2778

Qy       1164 CGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCCGGTTTATCATCCCCCTCGGGTGT 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2777 CGCTTCAAAAAATGGAACAACTTTACCGACCGCGCCCGGTTTATCATCCCCCTCGGGTGT 2718

Qy       1224 AATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTGGCAGATG 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2717 AATCAGAATAGCTGATGTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTGGCAGATG 2658

Qy       1284 GAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAGAGGGGAGCGCCACCAGAAGCAAT 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2657 GAACCTCTTGGCAACCGCTTCCCCGACTTCCTTAGAGAGGGGAGCGCCACCAGAAGCAAT 2598

Qy       1344 TTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCTTTTGGCGAAGAAGGAGAA 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2597 TTCGTGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCTTTTGGCGAAGAAGGAGAA 2538

Qy       1404 TAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAATCCTGAAGGCTCCTCAGAAACAG 1463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2537 TAGGGTTGGCACCAGCAGCGCACTTTGAATCTTGTAATCCTGAAGGCTCCTCAGAAACAG 2478

Qy       1464 CTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATATCCGAGTGT 1523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2477 CTCTTCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATATCCGAGTGT 2418

Qy       1524 AGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAATCGCAGTATCCGGAAT 1583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2417 AGTAAACATTCCAAAACCGTGATGGAATGGAACAACACTTAAAATCGCAGTATCCGGAAT 2358

Qy       1584 GATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAGAATCTCACGCAGGCAGTTCTATG 1643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2357 GATTTGATTGCCAAAAATAGGATCTCTGGCATGCGAGAATCTCACGCAGGCAGTTCTATG 2298

Qy       1644 AGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCAGAGGAGTTCATGATCAGTGCAAT 1703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2297 AGGCAGAGCGACACCTTTAGGCAGACCAGTAGATCCAGAGGAGTTCATGATCAGTGCAAT 2238

Qy       1704 TGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCGTATTCATTAAAACCGGGAGGTAG 1763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2237 TGTCTTGTCCCTATCGAAGGACTCTGGCACAAAATCGTATTCATTAAAACCGGGAGGTAG 2178

Qy       1764 ATGAGATGTGACGAACGTGTACATCGACTGAAATCCCTGGTAATCCGTTTTAGAATCCAT 1823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2177 ATGAGATGTGACGAACGTGTACATCGACTGAAATCCCTGGTAATCCGTTTTAGAATCCAT 2118

Qy       1824 GATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGCACGTTCAAAATTTTTTGCAACCC 1883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2117 GATAATAATTTTTTGGATGATTGGGAGCTTTTTTTGCACGTTCAAAATTTTTTGCAACCC 2058

Qy       1884 CTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATGCCCATACTGTTGAGCAATTCACG 1943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2057 CTTTTTGGAAACGAACACCACGGTAGGCTGCGAAATGCCCATACTGTTGAGCAATTCACG 1998

Qy       1944 TTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACTCCGATAAATAACGCGCCCAACAC 2003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1997 TTCATTATAAATGTCGTTCGCGGGCGCAACTGCAACTCCGATAAATAACGCGCCCAACAC 1938

Qy       2004 CGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACGACGATTCTGTGATTTGTATTCAG 2063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1937 CGGCATAAAGAATTGAAGAGAGTTTTCACTGCATACGACGATTCTGTGATTTGTATTCAG 1878

Qy       2064 CCCATATCGTTTCATAGCTTCTGCCAACCGAACGGACATTTCGAAGTACTCAGCGTAAGT 2123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 CCCATATCGTTTCATAGCTTCTGCCAACCGAACGGACATTTCGAAGTACTCAGCGTAAGT 1818

Qy       2124 GATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGCGTATCT 2183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1817 GATGTCCACCTCGATATGTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGCGTATCT 1758

Qy       2184 CTTCATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCATCTTCCAGCGGATAGAATGGCGC 2243
```

RESULT 10
US-09-743-194-19/c
; Sequence 19, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Belfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 12844
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-743-194-19

Query Match   59.6%; Score 1651.4; DB 4; Length 12844;
Best Local Similarity   99.9%; Pred. No. 2.2e-302;
Matches 1652; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   624 TTAAACGGCGATCTTTCCGCCCTTCTTCTTGGCCTTTATGAGGATTCTCTGATTTTTCTTGC 683
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2162 TTACACGGCGATCTTTCCGCCCTTCTTCTTGGCCTTTATGAGGATTCTCTGATTTTTCTTGC 2103

Qy   684 GTCGAGTTTCCGGTAAGACCTTTCGGTACTTCGTTCACAAAACACAACTCCTCCGGCAA 743
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2102 GTCGAGTTTCCGGTAAGACCTTTCGGTACTTCGTTCACAAAACACAACTCCTCCGGCAA 2043

Qy   744 CTTTTTCGCGGTTGTTACTTGACTCGGCGAGGCTAATCCACCGATCTCTTTTTTCCGTCATCGT 803
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2042 CTTTTTCGCGGTTGTTACTTGACTCGGCGAGGCTAATCCACCGATCTCTTTTTTCCGTCATCGT 1983

Qy   804 CTTTTCCGTCGTCCAAAAACAACAACCGGCAGAGTTCACCGGCGTCGGGGAAG 863
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1982 CTTTTCCGTCGTCCAAAAACAACAACCGGCAGAGTTCACCGGCGTCGGGGAAG 1923

Qy   864 ACCTGCGACACCTGGCTGCGTCGGAAGATGTTGGGGGTGTTGGAGCAAGATGGATTCCAATTCAGC 923
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1922 ACCTGCGACACCTGGCTGCGTCGGAAGATGTTGGGGGTGTTGGAGCAAGATGGATTCCAATTCAGC 1863

Qy   924 GGGAGCCACCTGATAGCCCTTTGTACTTAATCAGAGACTTCAGGCGGGTCAACGATGAAGAA 983
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1862 GGGAGCCACCTGATAGCCCTTTGTACTTAATCAGAGACTTCAGGCGGGTCAACGATGAAGAA 1803

Qy   984 GTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAAATGTAGCCATCCAACGATCCATCCTTGTCAAT 1043
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1802 GTGTTCGTCTTCGTCCCAGTAAGCTATGTCTCCAGAAATGTAGCCATCCAACGATCCATCCTTGTCAAT 1743

Qy   1044 CAAGGCGGTTGGTCGCTTCCGGATTGTTTACATAACCGGACTAATGACTTCTCAC 1103
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1742 CAAGGCGGTTGGTCGCTTCCGGATTGTTTACATAACCGGACATTAATCATGAGGACTCTCAC 1683

Qy   1104 ACACAGTTGCCCTCTTTTGATTAACGCCCAGGCGTTTTTCCGGTATCCCAGATCCACAACTT 1163
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1682 ACACAGTTCCCTCTTTTGATTAACGCCCAGCGTTTTCCCATGTTGCTCTCCATCTTCCAGCGC 1623

Qy   1164 CGCTTCAAAAAATGGAACAACTTTTACCGACCGGCGCGCCCGGGTTATCCCCTGGGGTGT 1223
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1622 CGCTTCAAAAAATGGAACAACTTTTACCGACCGGCGCGCCCGGGTTATCCCCTGGGGTGT 1563

Qy   1224 AATCAGAATAGCTGATGTAGTCTCAGTGACCCATATCCTTGCCTGATGATACCTGGCGAGATG 1283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1622 CGCTTCAAAAAATGGAACAACTTTTACCGACCGGCGCGCCCGGGTTATCCCCTGGGGTGT 1283

Db   1757 CTTCAATAGCCTTTATGCAGTTGCTCTCTCCAGCCGTTCCATCCTTCCAGCGGGATAGAAATGGGCGC 1698
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2244 CGGGGCCTTTCTTTCTTATGTTTTTTTGGCGGTTCTTCCATGG 2278
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1697 CGGGGCCTTTCTTTATGTTTTTTGGCGGTTCTTCCATGG 1663

Qy   1344 TTCGTGTAAATTAGATAATAAATCGTATTTGTTGTCAATCAGAGTTGCTTTTGGCGAAGAAGAA 1403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1284 GAACCTCTTGGCAACCGCTTCCCGACTTCCTTAGAGGAGGAGGCCCCACCAGAACCAAT 1343

Qy   1562 AATCAGAATAGCTGATGTAGTCTCAGTGACCCATATCCTTGCCTGATGATACCTGGCGAGATG 1503
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1502 GAAACCTCTTGGCAACCGCTTCCCGACTTCCTTAGAGGAGGAGGCCCCACCAGAACCAAT 1443

Qy   1404 TAGGGTTGGGCACCAGCGGCCGCACCTTTGAATCTTGTAATCAGAGTGCTTTTGGCGAAAACAG 1463
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1382 TAGGGTTGGGCACCAGCGGCCGCACTTTGAATCTTGTAATCAGAGTGCTTTTGGCGAAAACAG 1323

Qy   1764 ATGAGAGTGACGAAGTGTACATGACTGACTGAAATCCCTGGTAATCGTATTCATTGAAATCCAT 1823
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1082 TGTCTTTGTCCCTATCGTTCATAGCTTCTGCCAACCGAAGTGTTTCACTCGAAGTGTACTCAGT 1023

Qy   1704 TGTCTTGTCCCTATCGTTCATAGCTTCTGCCAACCGAAGTGTTTCACTCGAAGTGTACTCAGT 1763
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1142 AGGCAGAGCGGACACCTTTAGGCCAGACCAGTAGAATCCAGAGGAGTTCATGATCAGTGCAAT 1083

Qy   1644 AGGCAGAGCGGACACCTTTAGGCCAGACCAGTAGAATCCAGAGGAGTTCATGATCAGTGCAAT 1703
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1202 GATTTGATTGCCAAAAAATAGGAATCTTCTGGCATGCGAGAAATCTCACGCACGGCCAGTTCTATG 1143

Qy   1584 GATTTGATTGCCAAAAAATAGGAATCTTCTGGCATGCGAGAAATCTCACGCACGGCCAGTTCTATG 1643
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1262 AGTAAACATTCCAAAAACCGTGAATGGAACACACTTAAAATCGCAGTATCCGGAAT 1203

Qy   1524 AGTAAACATTCCAAAAACCGTGAATGGAACACACTTAAAATCGCAGTATCCGGAAT 1583
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1322 CTCTTCTTCAAATCTTATACATTGAAATCCAACATATCAAAATATTCCGAGTGT 1263

Qy   1464 CTCTTCTTCAAATCTTATACATTGAAATCCAACATATCAAAATATTCCGAGTGT 1523
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2102 GTCCGAGTTTCCGGTAAGACCTTTCGGTACTTCGTTCACAAAACACAACTCCTCCGGCAA 2043

Qy   1824 GATAATAAATTTTTTTGGATGAATTGGGAGCTTTTTTTTTTTGCACGTT 1883
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   962 GATAATAAATTTTTTTGGATGAATTGGGAGCTTTTTTTTTTTGCACGTT 903

Qy   1884 CTTTTTGGAAAACGAACAACCACGGTAGGCTGCGAAAAATGCCCATACTGTTGAGCAATTCACG 1943
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   962 GATAATAAATTTTTTTGGATGAATTGGGAGCTTTTTTTTTTTGCACGTT 903

Qy   1944 TTCATTATAAAAGTGGTTCGGCGGCGGCGCGCCAACTCGATAAAATAACGCCGGTTCAGC 2003
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   842 TTCATTATAAATGTCGTTCGGCGGCGCAACTGCAACTCCGGATAAAATAACGCCGGTTCAGCAC 783

Qy   2004 CGGCATAAAGAAATTGAAGACGAGTTTTCACTGCATAGCGACCATTCTGTTGATTTGTATTCAG 2063
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   782 CGGCATAAAGAAATTGAAGAGAGTTTTCACTGCATAGCGACCATTCTGTTGATTTGTATTCAG 723

Qy   2064 CCCATATCGTTTCATTAGCTTCTCTGCCAACCGAACCGGACACATTTCGAAGTACTCAGGTAAGT 2123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   722 CCCATATCGTTTCATTAGCTTCATTGCCAACCGAACGGACCATTTCGAAGTAGCTAAGT 663

Qy   2124 GATGTCCACCTCGATATGTTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGGGTATCT 2183
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   662 GATGTCCACCTCGATATGTTGCATCTGTAAAAGCAATTGTTCCAGGAACCAGGGGGTATCT 603

Qy   2184 CTTCATAGCCTTATGCAGTTGCTCTCTCCAGCCGGTTCCATCTTCCAGCGCGGATAGAATGGGCGC 2243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   602 CTTCAATAGCCTTATGCAGTTGCTCTCTCCATCTTCCAGCGCGGATAGAATGGGCGC 543

Qy   2244 CGGGGCCTTTCTTTATGTTTTTTGGCGGTCTTCCAT 2276
           ||||||||||||||||||||||||||||||||||||
Db   542 CGGGGCCTTTCTTTATGTTTTTTGGCGGTCTTCCAT 510

RESULT 11

US-09-743-194-20/c
Sequence 20, Application US/09743194

Patent No. 6716601

; GENERAL INFORMATION:
; APPLICANT: Belfield, Graham
APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; CURRENT FILING DATE: 2001-01-08
; CURRENT APPLICATION NUMBER: 2001-01-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
US-09-743-194-20
; SEQ ID NO 20
; LENGTH: 13073
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae

Query Match 59.6%; Score 1651.4; DB 4; Length 13073;
Best Local Similarity 99.0%; Pred. No. 2.2e-302;
Matches 1661; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

[DNA sequence alignment between Query (Qy) and Database (Db), with matched base pairs shown by vertical bars. The alignment spans positions 624–2391 on the query and corresponding positions on the database sequence.]

RESULT 12
US-08-354-240A-3/c
Sequence 3, Application US/08354240A

Patent No. 5670356
; GENERAL INFORMATION:
; APPLICANT: Sherf, Bruce A.
APPLICANT: Wood, Keith V.
; TITLE OF INVENTION: MODIFIED LUCIFERASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

```
;   ADDRESSEE: DeWitt Ross & Stevens, S.C.
;   STREET: 8000 Excelsior Drive, Suite 401
;   CITY: Madison
;   STATE: WI
;   COUNTRY: USA
;   ZIP: 53717-1914
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/354,240A
;   FILING DATE: 12-DEC-1994
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sara, Charles S.
;   REGISTRATION NUMBER: 30,492
;   REFERENCE/DOCKET NUMBER: 34506.029
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 608-831-2100
;   TELEFAX: 608-831-2106
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1650 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FRAGMENT TYPE: C-terminal
;   ORIGINAL SOURCE:
;   ORGANISM: Photinus pyralis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1651
;   US-08-354-240A-3

Query Match             59.5%;  Score 1649;  DB 1;  Length 1650;
Best Local Similarity   100.0%;  Pred. No. 4e-302;
Matches 1649;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy       628 ACGGCGGATCTTTCGCCCCTTCCGCCCCTTTTATGAGGATCTCTCTTGATTTTCTTGCGTCG    687
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1649 ACGGCGGATCTTTCGCCCCTTCCGCCCCTTTTATGAGGATCTCTCTTGATTTTCTTGCGTCG   1590

Qy       688 AGTTTTCCGGTAAGACCTTTCGGTACTTCGTCTCCACAAACACAACTCCTCCGCCGCAACTTT    747
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1589 AGTTTTCCGGTAAGACCTTTCGGTACTTCGTCTCCACAAACACAACTCCTCCGCCGCAACTTT   1530

Qy       748 TTCGCGGGTTGTTACTTGACTGGCGACGTAATCCAGATTCCTTTTCCGTCGTCATCGTCTTT    807
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1529 TTCGCGGGTTGTTACTTGACTGGCGACGTAATCCAGATTCCTTTTCCGTCGTCATCGTCTTT   1470

Qy       808 CCGTGCTCAAAACAACACAACGGGCGGCGGCGGCGGCGAAGTTCACCGTCATCGTCGGGAAGACCT    867
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1469 CCGTGCTCAAAACAACACAACGGGCGGCGGCGGCGGCGAAGTTCACCGTCATCGTCGGGAAGACCT   1410

Qy       868 GCGACACCCTGCGTCCGAAGAGTGGGGTGTTGGGACGCAAGATGGATTCCAATTCAGCGGGA    927
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1409 GCGACACCCTGCGTCCGAAGAGTGGGGTGTTGGGACGCAAGATGGATTCCAATTCAGCGGGA    927

Qy       928 GCCACCTGATAGCCCTTTGTACTTAATCAGAGACTTCAGCGCGGTCAACGATGAAGAAGTGT    987
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       928 GCCACCTGATAGCCCTTTGTACTTAATCAGAGACTTCAGCGCGGTCAACGATGAAGAAGTGT    987

Qy       988 TCGTCTTCTCCAGTAAGCTAGCCATCATCTCTCAGAGAATGTAGCCATCCATCCTTGTCAATCAAG   1047
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1289 TCGTCTTCTCCAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAG   1047

Qy      1048 GCGGTTGGGGTCGCCTTCCGGGATTGTTTTTACATAACCGGACACATATATAGGACCTTCACACAC   1107
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1229 GCGGTTGGGGTCGCCTTCCGGGATTGTTTTTACATAACCGGACACATATATAGGACCTTCACACAC   1170

Qy      1108 AGTTGCCTCTTTGATTAAGGCCCACACGGTTTCCCGGTATCCGAGATCCACAACCTTCGCT   1167
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1169 AGTTCGCCTCTTTGATTAAGGCCCACACGGTTTCCCGGTATCCGAGATCCACAACCTTCGCT   1110

Qy      1168 TCAAAAAATGGAACAACTTTACCGACGCGACCGCGCGCCCGGCGTTTTCCCGGGTATCATC   1227
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1109 TCAAAAAATGGAACAACTTTACCGACGCGACCGCGCGCCCGGCGTTTATCATCCCCGGGTGTTAATC   1050

Qy      1228 AGAATAGCTGATGTAGTCTCCAGTGAGCCCATAATCCTTGACCCATATCCGTGTAATGGAAC    1287
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1049 AGAATAGCTGATGTAGTCTCCAGTGAGCCCATAATCCTTGACCCATATCCGTGTAATGGAAC    990

Qy      1288 CTCTTGGCAACGCGCTTCCCCGACTTCCTTAGAGAGGGGAGGCCACCAGAGGCAATTTCG    1347
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       989 CTCTTGGCAACGCGCTTCCCCGACTTCCTTAGAGAGGGGAGGCCACCAGAGGCAATTTCG    930

Qy      1348 TGTAAATTAGATAAATCGTATTTGTCAATCAGAGTGCTTTTGCGGCTGAAGAAGTGAGTAGTA    1407
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       929 TGTAAATTAGATAAATCGTATTTGTCAAATCAGAGTGCTTTTTGGGCTGAAGAAGAGAAATAGG    870

Qy      1408 GTTGGCACCAGCGGCACTTTGAATCTTGTAATCCTGAAGGCTCCTCAGAAAACAGCTCT    1467
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       869 GTTGGCACCAGCGCACTTTGAATCTTGTAATCCTGAAGGCTCCTCAGAAAACAGCTCT    810

Qy      1468 TCTTCAAAATCTATACATTAAGACGACTCGAAATCCAACATCACATATTCAAATATCCGAGTGTAGTA    1527
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       809 TCTTCAAAATCTATACATTAAGACGACTCGAAATCCAACATCACATATTCAAAATATCCGGAGTAGTA    750

Qy      1528 AACATTCCAAAACCCGTGATGGAAATGGACAACACTTAAAAATCCGAGTATCCGGAATGATT    1587
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       749 AACATTCCAAAACCCGTGATGGAAATGGACAACACTTAAAAATCGACAGTATCCGGAATGATT    690

Qy      1588 TGATTGCCAAAAAATAGGATCTCTGCCATCGGAGAATCTCACGCAGCCAGTTCTATGAGGC    1647
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       689 TGATTGCCAAAAAATAGGATCTCTGCCATCGGAGAATCTCACGCAGCCAGTTCTATGAGGC    630

Qy      1648 AGAGCGGACACCTTTAGGCAGACCAGTAGATCCAGAGGAGTTCATGATCAGTGCAATTGTC    1707
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       629 AGAGCGGACACCTTTAGGCAGACCAGTAGATCCAGAGGAGTTCATGATCAGTGCAATTGTC    570

Qy      1708 TTGTCCCTATCGAAGGACTCTGGCACAAAATCGTATTCATTAAAACCCGGGAGGTAGATGA    1767
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       569 TTGTCCCTATCGAAGGACTCTGGCACAAAATCGTATTCATTAAAACCCGGGAGGTAGATGA    510

Qy      1768 GATGTGACGAACGTGTACATCGACTGAAATCCCTGGTAATCGTTTAGAATCCATGATA    1827
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       509 GATGTGACGAACGTGTACATCGACTGAAATCCCTGGTAATCGTTTAGAATCCATGATA    450

Qy      1828 ATAATTTTTGGATGAGTTGGGAGCTTTTTTGCAGTTCAAAAATTTTTGCAAACCCTTT    1887
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       449 ATAATTTTTTGGATGAGTTGGGAGCTTTTTTGCAGTTCAAAAATTTTTGCAAACCCTTT    390

Qy      1888 TTGGAAACGAACACCACGGTAGGCTGGGAAATGCCATACTGTTTGAGCAATTCACGTTCA    1947
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       389 TTGGAAACGAACACCACGGTAGGCTGGGAAATGCCATACTGTTTGAGCAATTCACGTTCA    330

Qy      1948 TTTATAAATGTCGTTCGCGCGGCGCGGCCAACTCCGATAAAATAACGGCGCCCAACGGC    2007
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       329 TTTATAAATGTCGTTCGCGCGGCGCGGCCAACTCCGATAAAATAACGGCGCCCAACGGC    270

Qy      2008 ATAAAGAAATTGAAGAGAGTTTCACTGCATACGACGAGATTCGTTGGATTTGTATTGGAGCCCA    2067
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       269 ATAAAGAAATTGAAGAGAGTTTTCACTGCATACGACGAGATTCGTTGATTTGTATTAGCCCA    210

Qy      2068 TATCGTTTCATAGCTTCTGCCAACCGAACGGACATTTCGAAGTACTCACGGTAAGTGATG    2127
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       209 TATCGTTTCATAGCTTCTGCCAACCGAACGGACATTTCGAAGTACTCACGGTAAGTGATG    150

Qy      2128 TCCACCTCGATATGTGCATCTGTAAAAGCAAATTGTTCCAGGAACCAGGGCGTATCTCTTC    2187
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       149 TCCACCTCGATATGTGCATCTGTAAAAGCAAATTGTTCCAGGAACCAGGGCGTATCTCTTC    90

Qy      2188 ATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCATCTTCAGCGGGATGAAATGGGCGGGGG    2247
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        90 ATAGCCTTATGCAGTTGCTCTCCAGCGGTTCCATCTTCAGCGGGATGAAATGGGCGGGGG    30
```

```
Db      89 ATAGCCTTATGCAGTTGCTCTCCAGCCGGATTGTCCATCTTCCAGCGGGATAGAATGGCGCCCGG      30
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    2248 CCTTTCTTTATGTTTTTGGCGTCTTCCAT 2276
           |||||||||||||||||||||||||||||
Db      29 CCTTTCTTTATGTTTTTGGCGTCTTCCAT 1
```

RESULT 13
US-09-602-628-3/c:
Sequence 3, Application US/09602628
Patent No. 6495355

GENERAL INFORMATION:

; APPLICANT:
; APPLICANT: Eames, Brian
; APPLICANT: Contag, Christopher
; TITLE OF INVENTION: Red-Shifted Luciferase
; FILE REFERENCE: SUN-127
; CURRENT APPLICATION NUMBER: US/09/602,628
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/140,598
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 3
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: red-shifted luciferase mutation
; NAME/KEY: CDS
; LOCATION: (1)...(1650)
; OTHER INFORMATION: Coding sequence of Clone 1
US-09-602-628-3

Query Match   59.5%; Score 1647.6; DB 4; Length 1686;
Best Local Similarity   98.6%; Pred. No. 7.4e-302;
Matches 1662; Conservative   0; Mismatches   24; Indels   0; Gaps   0;

```
Qy     591 TGGCCTAAGAGGCCGGAGTGTTACCCCAACCTTTAAACGGGACGATCTTTCCGCGCCTTCTT    650
           |   ||  ||  |||||||||
Db    1686 TGCTCGAACGCGGCCGCCGCCCGACTCTAGAATTACACGGGCGGATCTTTCCGCGCCCTTCTT    1627

Qy     651 GGCCCTTTATGAGGAATCTCTCTGATTTTTCTTGGCTCGAGTTTCCGGGTAAGACCTTTCGG    710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1626 GGCCCTTTATGAGGAATCTCTCTGATTTTTCTTGGCTCGAGTTTCCGGGTAAGACCTTTCG    1567

Qy     711 TACTTCGTCCACAAACACAACTCCTCCGGGCAACTTTGTGGTTGTTACTTGACTGGC    770
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1566 TACTTCGTCCACAAACACAACTCCTCCGGGTGTTAACTTTGACTGGC    1507

Qy     771 GACGTAATCCACGATCTCTCTTTTTCCGTCATCGTCTTTCCGTCTCCAAAACACACAGC    830
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1506 GACGTAATCCACGATCTCTCTTTTTCCGTGTCTTCAAAATTTTGCAAACACAACCGGC    1447

Qy     831 GGCGGGAAGTTCACCGGGCGGTCATCGTCGGGAGACAACCTGGCGGAAGTGTT    890
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1446 GGCGGGAGTTCACCGGGCGTCATGTGCGGGACACCTGGCGACACCTGGGC    1387

Qy     891 GGGGTGTTGGGACGAAGATGGAATTCAATTCAGCGGGAGCGAGCCACCTGATAGCCTTTGTACTT    950
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1386 GGGGTGTTGGGACGAAGATGGAATTCCAATTCAGCGGGAGCGAGCCACCTGATAGCCTTTGTACTT    1327

Qy     951 AATCAGAGACTTCAGCGGCGGTCACGATGAAGAAGAGTGTTCGTCTTCGTCTTCCAGTAAGCTAT   1010
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1326 AATCAGAGACTTCAGCGGGTCAACGACAGAAGTGTTCGTCTTCGTCGTCCCAGTAAGCTAT   1267

Qy    1011 GTCTCCGGAAATGTAGCCATCCATCCCTTGTCAATCAAACAAGCCGTTGGTTCGCTTCCCGGATTGTT   1070
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1266 GTCTCCCAGAATGTAGCCATCCATCCTTGTCAATCAAGGCGTTGGTTCGCTTCCCGGATTGTT   1207

Qy    1071 TACATAACCGGACAATAATCATTAGGACCTCTCACACAGTTCGGCCTTCTTCTTTGATTAACGCC   1130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1206 TACATAACCGGACAATAATCATTAGGACCCTCTCACACAGTTCGGCCTTCTTCTTTGATTAACGCC   1147
```

```
Qy    1131 CAGGCGGTTTTCCCGCGTATCCAGATCCGCATCCACAACACCTTGCTTCTTGCTTCAAAAAATGGAACAACTTTACC   1190
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1146 CAGGCGTTTTCCCGCGTATCCAGATCCCAGATCCACACAACACCTTGCTTCTTGCTTCAAAAAATGGAACAACTTTACC   1087

Qy    1191 GACCGCGCCCGGGTTTATCATCCCCTCGGGGTGTAATCAGAATAGCTGATCTCAGT   1250
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1086 GACCGCGCCCGGGTTTATCATCCCCTCGGGGTGTAATCAGAATAGCTGATCTCAGT   1027

Qy    1311 TTCCTTAGAGAGTGGAGGGGAGCGCGCCACCAGAAGCAGCAAATTTCGTGTAAAATTAGAATATCGTATTT   1370
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     966 TTCCTTAGAGAGTGGAGGGGAGCGCGCCACCAGAAGCAGCAAATTTCGTGTAAAATTAGAATATCGTATTT    907

Qy    1371 GTCAATCAGAGTGCTTTTGGCGAAGAGAAGGAGAATAGGGGTTGGCACCAGCAGCGCACTTTG   1430
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     906 GTCAATCAGAGTGCTTTTGGCGAAGAGAAGGAGAATAGGGGCACCAGCAGCGCACTTTG    847

Qy    1431 AATCTTGTAATCTGAAGGCTCCTCAGAAAACAGCTCTTCTTCAAATCTATACATTAAAGAC   1490
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     846 AATCTTGTAATCCTGAAGGCTCCTCAGAAAACAGCTCTTCTTCAAATCTATACATTAAAGAC    787

Qy    1491 GACTCGAAATCCACTATCATATCAAATATCCGAGTGGTAGTGTAGTAGTAAAACATTCCAAAAACCGTGATGGAA   1550
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     786 GACTCGAAATCCACATATCAAATATCCGAGTGGTAGTGTAAAACATTCCAAAAACCGTGATGGAA    727

Qy    1551 TGGAACAACACTTAAAATGCAGTATCCCGGAATGTCCGGAATCTCT   1610
           ||||||||||||||||||||||||||||||||||||||||
Db     726 TGGAACAACACTTAAAATGCAGTATCCCGGAATGATTTGGAATCTCT    667

Qy    1611 GGCATGCGAGAATCTCAGCGCAGGCAGTTCTATGAGGCGCAGAGCGACACCTTTAGGCAGACC   1670
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     666 GGCATGCGAGAATCTCACGCGCAGGCAGTTCTATGAGGCGCAGAGCGACACCTTTAGGCAGACC    607

Qy    1671 AGTAGGATCCAGAGGAGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGAGACTCTGG   1730
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     606 AGTAGGATCCAGAGGAGAGTTCATGATCAGTGCAATTGTCTTGTCCCTATCGAAGGAGACTCTGG    547

Qy    1731 CACAAAAATCGTATTCATTAAAACCGGAGGTAGATGAGTAGATGAGAAGTGTACATCGA   1790
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     546 CACAAAAATGTATTCATTAAAACCGGAGGTAGATGAGTAGATGAGAAGTGTACATCGA    487

Qy    1791 CTGAAAATCCCTGGTAAATCGGTTTAGAAATCCATGATAATAATTTTTTGGATGATTGGGGAG   1850
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     486 CTGAAAATCCCTGGTAAATCGGTTTAGAAATCCATGATAATAATTTTTTGGATGATTGGGGAG    427

Qy    1851 CTTTTTTTGCACGTTCAAAAATTTTTGGAAACGAACGACCACGGGTAGG   1910
           ||||||||||||||||||||||||||||||||||||||||||||
Db     426 CTTTTTTTGCACGTTCAAAAATTTTTGGAAACGAACGACCACGGGTAGG    367

Qy    1911 CTGCAAAATGCCCAACTGTTGACGCAATTCACGTTCATTATAAATGTCGTTCGCGGGGCGC   1970
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 CTGGGAAAATGCCCATACTGTTGACGCAATTCACGTTCATTATAAATGTCGTTCGCGGGGC    307

Qy    1971 AACTGCAACTCCGGATAAATAAACCGCCCAACACCGGGCCATAAAGAATTGAAGAGAGTTTTC   2030
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 AACTGCAACTCCGATAAATAAACCGCCCAACACCGGGCCATAAAGAATTGAAGAGAGTTTTC    247

Qy    2031 ACTGCATACGGACGATTCGTGTGATTTGTATTCAGCCGCCATATCGTTCATAGCTTCTGCCAA   2090
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     246 ACTGCATACGGACGATTCGTGTGATTTGTATTCAGCCGCCATATCGTTCATAGCTTCTGCCAA    187

Qy    2091 CCGAACGGACATTCGAGCTACTCAGCCGTAAGTGAGTGAGATGTCCACCTCGATAGTGCATCTGT   2150
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CCGAACGGACATTCGAGCTACTCAGCCGTAAGTGAGTGAGATGTCCACCTCGATAGTGCATCTGT    127

Qy    2151 AAAAGCAAATTGTTCCAGAGAACCAGGCGGGCGGTTATCTCTTCATTAGCCTTATCGCAGTTGTCTC   2210
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     126 AAAAGCAAATTGTTCCAGGAACCAGGGCGTATCTCTTCATTAGCCTTATCGCAGTTGTCTCTCC    67
```

```
Qy   2211  AGCGGGTTCCATCTTCCAGCGGGATAGAATGCGCCGGGGCCCGGGGCCCTTTCTTTATGTTTTTTGGCGGTC  2270
           ||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||
Db     66  AGCGGGTTCCATCTTCCAGCGGGATAGAATGGGCGCCGGGGCCCGGGGCCCTTTCTTTATGTTTTTTGGCGTC  7

Qy   2271  TTCCAT  2276
           ||||||
Db      6  TTCCAT  1
```

RESULT 14
US-08-354-240A-5/c
Sequence 5, Application US/0835240A
; Patent No. 5670356
; GENERAL INFORMATION:
;   APPLICANT: Sherf, Bruce A.
;   APPLICANT: Wood, Keith V.
;   TITLE OF INVENTION: MODIFIED LUCIFERASE
;   NUMBER OF SEQUENCES: 20
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: DeWitt Ross & Stevens, S.C.
;   STREET: 8000 Excelsior Drive, Suite 401
;   CITY: Madison
;   STATE: WI
;   COUNTRY: USA
;   ZIP: 53717-1914
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/354,240A
;   FILING DATE: 12-DEC-1994
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sara, Charles S.
;   REGISTRATION NUMBER: 30,492
;   REFERENCE/DOCKET NUMBER: 34506.029
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 608-831-2100
;   TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1653 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FRAGMENT TYPE: C-terminal
;   ORIGINAL SOURCE:
;   ORGANISM: Photinus pyralis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1650
US-08-354-240A-5

Query Match 59.3%; Score 1643; DB 1; Length 1653;
Best Local Similarity 100.0%; Pred. No. 5.5e-301;
Matches 1643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    628  ACGGCGGATCTTTCCGCCCTTTCTTGGCCTTTATGAGGATCTCTCTGATTTTTTCTTGCGTCG  687
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1652  ACGGCGGATCTTTCCGCCCTTTCTTGGCCTTTATGAGGATCTCTCTGATTTTTTCTTGCGTCG  1593

Qy    688  AGTTTTTCCGGTAAGACCTTTCGGTACTTCGTCTCCACAAAACAACTCCCGCGCGCAACTTT  747
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1592  AGTTTTTCCGGTAAGACCTTTCGGTACTTCGTCTCCACAAAACAACTCCCGCGCGCAACTTT  1533

Qy    748  TTCCGCGGTTGTTACTTGACTGGCGGGGACGTAATCACAGATCTCTTTTTTCCGTCTCTTT  807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1532  TTCGCGGTTGTTACTTGACTGGCGGGGACGTAATCACAGATCTCTTTTTTCCGTCTCTTT  1473

Qy    808  CCGTCTCTCCAAAACAACAACGGCGGGGGAAGTTCACCGCGCGTCATCGTCGGGAAGACCT  867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1472  CCGTGCTCTCCAAAACAACAACGGCGGGGGAAGTTCACCGCGCGTCATCGTCGGGAAGACCT  1413

Qy    868  GCGACACCTCCGTCGAAGAGTTTGGGGTGTTGGGGATTCCAATTCAGCGGGGA  927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1412  GCGACAACCTGCCTCGAAGAGTGTTGGGGTGTTGGGGATTCCAATTCAGCGGGGA  1353

Qy    928  GCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGGTCAACGATGAAGAAGTGT  987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1352  GCCACCTGATAGCCTTTGTACTTAATCAGAGACTTCAGGCGGGTCAACGATGAAGAAGTGT  1293

Qy    988  TCGTCTTCGTCGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAG  1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1292  TCGTCTTCGTCGTCCCAGTAAGCTATGTCTCCAGAATGTAGCCATCCATCCTTGTCAATCAAG  1233

Qy   1048  GCGGTTCGGTTCGGGGATTGTTTACATAACGGGACATAATCATAGGGACTTCTCACACAC  1107
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1232  GCGGTTCGGTTCGGGGATTGTTTACATAACGGGACATAATCATAGGGACTTCTCACAC  1173

Qy   1108  AGTTCGGCCTCTTTGATTAAGCCCAAGCGTTTTCCCGGTACCGTTTCCAAACCCTTTCGCT  1167
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1172  AGTTCGGCCTCTTTTGATTAAGCCCCAGCGTTTTCCCGGTACCGTTTCCCAAACCTTTCGCT  1113

Qy   1168  TCAAAAAATGGAACAACTTTACCGACCGCGCCCGGGTTTTATCATCCCCTCGGGGTGTAATC  1227
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1112  TCAAAAAATGGAACAACTTTACCGACCGCGCCCGGGTTTTATCATCCCCTCGGGGTGTAATC  1053

Qy   1228  AGAATAGCTGATGTGATTGTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTTGGCGGGAC  1287
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1052  AGAATAGCTGATGTGATTGTAGTCTCAGTGAGCCCATATCCTTGCCTGATACCTTGGCGGGAC  993

Qy   1288  CTCTTGGCAACCGCTTCCCCGGACTTCCTTAGAGAGGGGGACGCGCCACCAGAAGCAATTTCG  1347
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    992  CTCTTGGCAACCGCTTCCCCGGACTTCCTTAGAGAGGGGGACGCGCCACCAGAAGCAATTCG  933

Qy   1348  TGTAAATTAGATAAAATCGTATTTGTCAATCAGAGTGCTTTTGGGGAGGAAGAGGAATAGG  1407
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    932  TGTAAATTAGATAAAATCGTATTTGTCAATCAGAGTGCTTTTGGGGAGGAAGAGGAATAGG  873

Qy   1408  GTTGGCACCAGCAGCGCCACTTTGAATCTTGTAAATCCTGAAGGCTCCTCAGAAAACAGCTCT  1467
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    872  GTTGGCACCAGCAGCGCCACTTTGAATCTTGTAAATCCTGAAGGCTCCTCAGAAAACAGCTCT  813

Qy   1468  TCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATATCGAGTGTAGTA  1527
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    812  TCTTCAAATCTATACATTAAGACGACTCGAAATCCACATATCAAATATCGAGTGTAGTA  753

Qy   1528  AACAATTCCAAAACCCGTAGTGGAATGGAACAACACTTAAAATGCAGTATCCGGAATGATT  1587
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    752  AACAATTCCAAAACCCGTAGTGATGGAATGGAACAACACTTAAAATGCAGTATCCGGAATGATT  693

Qy   1588  TGATTGCCAAAAAATAGGATCTCTGGCATGCGGAGAATCCACGCAGGCAGTTCTATGAGGC  1647
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    692  TGATTGCCAAAAAATAGGATCTCTGGCATGCGGAGAATCCACGCAGGCAGTTCTATGAGGC  633

Qy   1648  AGAGCGGCACCTTTAGGCGACCAGTAGATCCAGAGGAGTTCATATCACTGCAATTGTC  1707
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    632  AGAGCGGCACCTTTAGGCGACCAGTAGATCCAGAGGAGTTCATGATCACTGCAATTGTC  573

Qy   1708  TTGTCCCTATCGAAGGACTCTGGCACAAAAATCGTATTCATTAAAACCGGAGGTAGATGA  1767
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    572  TTGTCCCTATCGAAGGACTCTGGCACAAAAATCGTATTCATTAAAACCGGAGGTAGATGA  513

Qy   1768  GATGTGACGAACGTGTACATCGACTGAAATCCCGGTAATCGTTTAGAATCCATGATA  1827
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    512  GATGTGACGAACGTGTACATCGACTGAAATCCCGGTAATCGTTTAGAATCCATGATA  453

Qy   1828  ATAATTTTTTGGATGATTGGGGAGCTTTTTTTGCACGTTCAAAAATTTTTGCAACCCCTTT  1887
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    452  ATAATTTTTTGGATGATTGGGGAGCTTTTTTTGCACGTTCAAAAATTTTTGCAACCCCTTT  393

Qy   1888  TTGGAAACGAACACCACGGTAGGCTCGGAAATGCCCATACTGTTGAGCAATTCACGTTCA  1947
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 15
US-09-602-628-1/c
Sequence 1, Application US/09602628
Patent No. 6495355

GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: Contag, Christopher
; APPLICANT: Eames, Brian
; TITLE OF INVENTION: Red-Shifted Luciferase
; FILE REFERENCE: SUN-127
; CURRENT APPLICATION NUMBER: US/09/602,628
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/140,598
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 4.0

SEQ ID NO 1
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Red shifted luciferase mutation
; NAME/KEY: CDS
; LOCATION: (1)...(1650)
US-09-602-628-1

Query Match 59.3%; Score 1642.8; DB 4; Length 1686;
Best Local Similarity 98.4%; Pred. No. 6e-301;
Matches 1659; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 392 TTGGAAACGAACACCACGGGTAGGCTGCGAAAATCCCATACTGTTGAGCAATTCACGTTCA 333

Qy 591 TGGCCTAAGAGGCCGGAGTGTTTACCCAACCTTTAAACGGCGATCTTTTCCGCGCCTTCTT 650

Best Match text and sequence alignment blocks (Qy/Db paired rows) follow.

```
Db   366 CTGCGAAATGCCCATACTGTTTGAGCAAATTCACGTTCATTATTAAATGTCGTTCGCGGGGGGC 307

Qy  1971 AACTGCAACTCCGATAAAATAAACGCCCCAACACCCGCATAAAAGAAATTGAAGAGAGTTTTC 2030
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306 AACTGCAACTCCGATAAAATAAACGCCCCAACACCCGCATAAAAGAAATTGAAGAGAGTTTTC 247

Qy  2031 ACTGCATACGACGATTCTGTTGATTTGTATTCAGCCCATATCGTTTCATAGCCTTCTGCCAA 2090
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246 ACTGCATACGACGATTCTGTTGATTTGTATTCAGCCCATATCGTTTCATAGCCTTCTGCCAA 187

Qy  2091 CCGAACGGACATTTCGAAGTACTCAGCGTAAGTGAATGTCCACCTCGATATGTGCAATCTGT 2150
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186 CCGAACGGACATTTCGAAGTACTCAGCGTAAGTGAATGTCCACCTCGATATGTGCAATCTGT 127

Qy  2151 AAAAGCAATTGTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCTCC 2210
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126 AAAAGCAATTGTTCCAGGAACCAGGGCGTATCTCTTCATAGCCTTATGCAGTTGCTCTCTCC 67

Qy  2211 AGCCGGTTCCATCTTCCAGCGGGGATAGAAATGGCGCCGGGGCCTTTCTTTATGTTTTTGGCGGTC 2270
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66 AGCCGGTTCCATCTTCCAGCGGGGATAGAAATGGCGCCGGGGCCTTTCTTTATGTTTTTGGCGGTC 7

Qy  2271 TTCCAT 2276
         ||||||
Db     6 TTCCAT 1
```

Search completed: September 29, 2005, 15:08:15
Job time : 460 secs